TAMPEREEN YLIOPISTO

Valinnaisten opintojen syventäviin opintoihin kuuluva tutkielma

Lauri Kumpulainen

Büchin automaateista

Luonnontieteiden tiedekunta
Tietojenkäsittelytieteiden tutkinto-ohjelma

Huhtikuu 2017

Tampereen yliopisto

Luonnontieteiden tiedekunta

KUMPULAINEN, LAURI: Büchin automaateista

Valinnaisten opintojen syventäviin opintoihin kuuluva tutkielma, 41 s.

Tietojenkäsittelytieteiden tutkinto-ohjelma

Huhtikuu 2017

Tiivistelmä

Tutkielma käsittelee Büchin automaatteja. Ensiksi esitetään aakkostot, merkkijonot ja kielet, jonka jälkeen edetään deterministiseen ja epädeterministiseen äärelliseen automaattin. Osoitetaan, että näiden ilmaisuvoima on sama, ja todistetaan Kleenen lause äärelliselle automaatille. Sen jälkeen esitetään äärettömät merkkijonot, ω-säännölliset kielet, deterministien Büchin automaatti ja epädeterministinen Büchin automaatti. Lopuksi osoitetaan Kleenen lause Büchin automaatille ja että deterministisen Büchin automaatin ilmaisuvoima on aidosti heikompi kuin epädeterministisen Büchin automaatin.

1 Johdanto

Automaation idea juontaa juurensa aina tieteellisen vallankumouksen aikaan 1600-luvulle, jolloin alati kasvanut tietämys universumista synnytti ajatuksen, että universumi olisi vain valtava kone. Myös R. Descartes (1596-1650) pohti mahdollisuutta, että kaiken elävän käyttäytyminen voitaisiin kuvailla matematiikan keinoin. Ihmisen käyttäytymisen suhteen väittely determinismin ja vapaan tahdon suhteen raivosi vuosisatoja ja jatkunee vielä tulevaisuudessakin.

Ihmiskunnan teollistuessa ja mekaniikan kehittyessä myös matemaatikot alkoivat pohtimaan, voisiko matemaattisia väitteitä todistaa mekaanisesti jollain tavalla. Toisin sanoen etsittiin algoritmia, jolla todistaa matemaattiset ongelmat. Mutta 1900-luvun alusta alkoi näyttämään, ettei se välttämättä ole mahdollista. Lopulta vuonna 1931 K. Gödel (1906-1978) todisti kuuluisat epätäydellisyyslauseensa, jotka osoittavat lukuteorian sisältävän aksiomaattisen järjestelmän olevan epätäydellinen.

Tämä ei kuitenkaan ollut vahingollista matematiikan ja tietojenkäsittelytieteen suhteen, sillä näiden tapahtumien myötä syntyi täsmällinen määritelmä algoritmille. Syntyi kaksi uutta tieteenalaa, jotka ovat laskettavuuden teoria ja formaalien kielten ja automaattien teoria. Näistä jälkimmäisessä yksinkertaisimmat algoritmit ovat niitä, joita voidaan soveltaa äärellisillä automaateilla.

Lopulta 1950-luvulla Stephen Kleene (1909-1994) tutki ensimmäisenä äärellisiä automaatteja ja löysi niille monia tärkeitä sovelluksia tietojenkäsittelytieteessä. Seuraavina vuosikymmeninä matemaatikot S. Eilenberg (1913-1998), M.-P. Schützenberger (1920-1996) ja J. Rhodes (1937-) kehittivät äärellisten automaattien matematiikkaa. Aivan viime aikoina äärellisiä automaatteja on sovellettu esimerkiksi kombinatorisessa ryhmäteoriassa ja symbolisessa dynamiikassa.

Samoihin aikoihin 1960-luvulla sveitsiläinen loogikko ja matemaatikko J. Büchi (1924-1984) kehitti äärellisen automaatin muunnoksen, joka kykenee tunnistamaan äärettömiä merkkijonoja. Tätä automaattia kutsutaan Büchin automaatiksi.

Tutkielman toisessa luvussa esitetään aakkostot ja merkkijonot, josta siirrytään kieliin, jonka jälkeen esitellään ensin deterministinen automaatti ja sitten epädeterministinen automaatti. Lisäksi osoitetaan, että näiden ilmaisuvoima on sama. Lopuksi esitetään säännölliset lausekkeet ja todistetaan Kleenen lause äärellisille automaateille.

Tutkielman kolmannessa luvussa esitetään vastaavasti äärettömät merkkijonot ja ω-säännölliset kielet. Epädeterministinen Büchin automaatti esitellään ennen determinististä Büchin automaattia. Lopuksi osoitetaan Kleenen lause Büchin automaateille ja että deterministinen Büchin automaatti on ilmaisuvoimaltaan aidosti heikompi kuin epädeterministinen Büchin automaatti.

Eräs mielenkiintoinen äärellisten merkkijonojen ja Büchin automaattien sovellus on LTL-logiikan kielellä määriteltyjen väitteiden formaali verifiointi Büchin automaattien avulla.

Rozier kirjoittaa [5] LTL-logiikan symbolisesta mallintarkastuksesta Büchin automaattien avulla. Hän käyttää esimerkkisovelluksena eräänlaista automaattista ilmatilan valvontaa lennonjohdossa.

Tutkittavasta ongelmasta muodostetaan malli, jonka jälkeen muodostetaan LTL-logiikan kaava, joka ilmaisee mallilta vaadittavaa vaatimusta. Tämän jälkeen tarkistetaan, että onko tämä LTL-logiikan kaava toteutuva tässä mallissa.

Tarkistus onnistuu siten, että muodostetaan edellä mainitun LTL-logiikan kaavan negaatiosta Büchin automaatti. Ääretön merkkijono muodostetaan tällöin aakkostosta, joka on muodostettu LTL-logiikan propositiosymboleiden potenssijoukosta. Sen jälkeen tutkitaan, onko olemassa polkua, jonka tämä LTL-logiikan kaavan negaatiosta muodostettu Büchin automaatti hyväksyy. Jos sellainen polku löytyy, niin se toimii vastaesimerkkinä. Jos taas sellaista ei löydy, niin tällöin malli toteuttaa LTL-logiikan kaavan avulla ilmaistun vaatimuksen.

Käytännön sovelluksien kannalta on lohdullista, että on olemassa eri algoritmeihin pohjautuvia automaattisia mallintarkastustyökaluja. Nimittäin ongelmaksi käytännön sovelluksissa muodostuu yleisesti tilaräjähdys, jossa siis tilojen lukumäärä kasvaa eksponentiaalisesti.

2 Äärelliset automaatit

Tässä luvussa esitellään määritelmiä ja tuloksia äärellisistä automaateista pääasiassa lähteen [3] mukaan, joskin vertaillen lähteisiin [2] ja [1]. Tämän luvun tuloksien todistuksia on sivuutettu, ja luku toimiikin alustuksena ja vertailukohtana seuraavalle luvulle, joka käsittelee äärettömiä automaatteja.

2.1 Aakkostot ja merkkijonot

Määritelmä 2.1. *Aakkosto* A on äärellinen joukko, jonka alkiot ovat mitä tahansa symboleja. Aakkoston A symboleita kutsutaan *kirjaimiksi* tai *merkeiksi*. Aakkoston A merkeistä muodostettua äärellistä jonoa $(a_1, a_2, \ldots, a_n)$ kutsutaan *merkkijonoksi*. Merkkijonolle käytetään kuitenkin lyhyempää merkintää $a_1a_2 \cdots a_n$. Aakkoston A merkkien *lukumäärää* merkitään $|A|$. Sovitaan, että *tyhjää merkkijonoa* merkitään kreikkalaisella kirjaimella ε.

Määritellään, että A^* on aakkoston A *kaikkien merkkijonojen joukko*, ja että A^+ on aakkoston A *kaikkien merkkijonojen joukko lukuun ottamatta tyhjää merkkijonoa* ε.

Jos x on merkkijono, niin $|x|$ on merkkijonon x *pituus*. Jos a on aakkoston A merkki, niin $|x|_a$ ilmaisee merkin a esiintymien lukumäärän merkkijonossa x. Aakkoston A merkkijonojen x ja y sanotaan olevan *yhtäsuuret*, mikäli ne sisältävät samat merkit samassa järjestyksessä.

Esimerkki 2.1. Suomen kielen aakkosto lienee tämän tutkielman lukijoille tutuin, ja se voidaan ilmaista muodossa $\{a, b, \ldots, \text{å}, \text{ä}, \text{ö}\}$. Tietokoneet tulevat toimeen binääriaakkostolla $\{0, 1\}$.

Jos ääretön aakkosto sallittaisiin, niin predikaattilogiikan loogisen aakkoston taas voisi ilmaista muodossa $\{v_i \mid i \in \mathbb{N}\} \cup \{\neg, \wedge, \vee, \rightarrow, \leftrightarrow, \forall, \exists, =, (,)\}$.

Määritelmä 2.2. Olkoot $x, y \in A^*$. Merkkijonoista x ja y voidaan muodostaa uusi merkkijono $x \cdot y$, jota kutsutaan *merkkijonojen x ja y konkatenaatioksi*, liittämällä merkkijonot peräkkäin. Yleisesti merkitään vain xy.

Olkoot edelleen $x, y \in A^*$. *Konkatenaation pituus* määritellään asettamalla $|xy| = |x| + |y|$. Tyhjälle merkkijonolle ε pätee $\varepsilon x = x = x\varepsilon$ jokaisella $x \in A^*$.

Tyhjä merkkijono on tällöin konkatenaation suhteen *neutraalialkio*. Osoitetaan, että konkatenaatio on *liitännäinen*. Olkoot $x, y, z \in A^*$, missä $x = a_1a_2 \cdots a_n$, $y = b_1b_2 \cdots b_n$ ja $z = c_1c_2 \cdots c_n$. Nyt

$$\begin{aligned}(xy)z &= (a_1a_2 \cdots a_nb_1b_2 \cdots b_n)c_1c_2 \cdots c_n \\ &= a_1a_2 \cdots a_nb_1b_2 \cdots b_nc_1c_2 \cdots c_n \\ &= a_1a_2 \cdots a_n(b_1b_2 \cdots b_nc_1c_2 \cdots c_n) \\ &= x(yz).\end{aligned}$$

Konkatenaatio on siis aakkoston A kaikkien merkkijonojen joukossa A^* määritelty *binäärioperaatio*, joka osoitettiin edellä liitännäiseksi, jolloin tarkasteltava rakenne on *puoliryhmä*. Koska konkatenaatiolla on myös neutraalialkio, kuten niin ikään yllä osoitettiin, niin rakenne on *monoidi*. Konkatenaatiolla varustettua joukkoa A^+ kutsutaan myös *vapaaksi puoliryhmäksi* ja joukkoa A^* *vapaaksi monoidiksi*, koska kaikki niiden merkkijonot voidaan esittää yksikäsitteisesti aakkoston A merkkejä konkatenoimalla.

Esimerkki 2.2. Olkoot $A = \{a, b\}$ ja $x, y \in A^*$. Asetetaan $x = ab$ ja $y = ba$, jolloin $xy = abba$ ja $yx = baab$, joten $xy \neq yx$. Konkatenaatio ei siis ole vaihdannainen operaatio.

Määritelmä 2.3. Olkoot $x \in A^*$. Määritellään, että merkkijonon x *potenssi* on konkatenaatio itsensä kanssa n-kertaa asettamalla $x^n = \underbrace{x \cdots x}_{\text{n kpl}}$, missä $n \in \mathbb{N}$. Lisäksi määritellään $x^0 = \varepsilon$.

Olkoot $x \in A^*$ ja $m, n \in \mathbb{N}_0$. Edellisen määritelmän nojalla pätee selvästi, että $x^m x^n = x^{m+n}$.

Määritelmä 2.4. Olkoot $x, y, z \in A^*$ ja $u = xyz$. Tällöin y on merkkijonon u *tekijä*, x on *etuosa* ja z on *jälkiosa*. Mikäli ainakin toinen merkkijonoista x ja y on epätyhjä, on y merkkijonon u *aito tekijä*. Mikäli $x \neq u$, niin x on merkkijonon u *aito etuosa* ja edelleen, mikäli $z \neq u$, niin x on merkkijonon u *aito jälkiosa*. Merkkijono u on merkkijonon v *alimerkkijono*, jos $u = a_1 \dots a_n$, missä $a_i \in A$ ja on olemassa sellaiset merkkijonot $x_0, \dots, x_n$, että $v = x_0a_1x_1 \cdots x_{n-1}a_nx_n$. Olkoon $x \in A^*$. Tällöin merkkijonoa $x = u_1 \dots u_n$, missä $u_i \in A^*$, kutsutaan merkkijonon x *tekijöihin jaoksi*. Tekijöihin jaossa käytetään yleisesti merkkiä $\cdot$ korostamaan tekijöihin jakoa.

Esimerkki 2.3. Tarkastellaan aakkostoa $\{a, b\}$ ja sen merkkijonoa $x = baba$. Tällöin merkkijonon x etuosat ovat $\varepsilon, b, ba, bab, baba$, jälkiosat ovat $\varepsilon, a, ba, aba, baba$ ja tekijät ovat $\varepsilon, a, b, ba, ab, bab, aba, baba$. Merkkijonot bab ja ba ovat merkkijonon x alimerkkijonoja. Merkkijonon x eräs tekijöihin jako on $ba \cdot ba$.

Määritelmä 2.5. *Verkko* on pari $G = (V, E)$, joka koostuu äärellisestä joukosta V, jonka alkiot ovat *solmuja*, ja joukosta E, jonka alkiot ovat solmupareista muodostettuja *särmiä*.

Verkon polku on sellainen solmujen muodostama jono $v_1, v_2, \ldots, v_k$, $k \geq 1$, että on olemassa särmä (v_k, v_{k+1}) jokaisella i, $1 \leq i < k$. *Polun* $v_1, v_2, \ldots, v_k$ *pituus* on $k - 1$ ja, jos $v_1 = v_k$, on kyseessä *sykli*.

Määritelmä 2.6. *Suunnattu verkko* on niin ikään pari $G = (V, E)$, joka koostuu äärellisestä joukosta V, jonka alkiot ovat *solmuja* ja joukosta E, jonka alkiot ovat järjestetyistä solmupareista muodostettuja *nuolia*. Merkintä $v \rightarrow w$ tarkoittaa nuolta solmusta v solmuun w.

Suunnatun verkon polku on sellainen solmujen muodostama jono $v_1, v_2, \ldots, v_k$, $k \geq 1$, että on olemassa nuoli $v_i \rightarrow v_{i+1}$ jokaisella i, $1 \leq i < k$. Sanotaan, että $v_1, v_2, \ldots, v_k$ on polku solmusta v_1 solmuun v_k. Jos $v \rightarrow w$ on nuoli, niin sanotaan, että v on solmun w *edeltäjä* ja w on solmun v *seuraaja*.

Määritelmä 2.7. *Puu* on suunnattu verkko, jolla on seuraavat ominaisuudet:

i) On olemassa yksi yksikäsitteinen solmu, jota kutsutaan *juureksi*, jolla ei ole lainkaan edeltäjää.

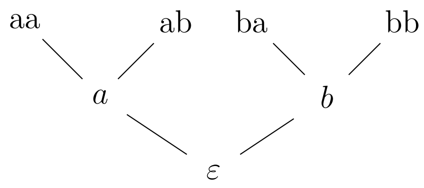
ii) Jokaisella solmulla, paitsi juurella, on ainoastaan yksi edeltäjä.

iii) Jokaisen solmun seuraajat järjestetään vasemmalta oikealle.

Solmua, jolla ei ole seuraajaa, kutsutaan *lehdeksi*.

Määritelmä 2.8. Olkoon A mielivaltainen aakkosto ja olkoot x, y sen merkkijonoja. Määritellään, että *puu yli joukon* A^* on sellainen puu, että sen *juuri* on tyhjä merkkijono ε, ja jonka solmut nimetään joukon A^* alkioiden mukaan. Mikäli w on puun mielivaltainen solmu, niin siitä kasvavat solmut ovat $wa_1, \ldots, wa_n$. *Puujärjestys yli joukon* A^* määritellään asettamalla $x \leq y$, jos ja vain jos $|x| < |y|$, tai $|x| = |y|$ ja merkkijono x on merkkijonon y vasemmalla puolen joukon A^* puussa.

Esimerkki 2.4. Olkoot $A = \{a, b\}$ ja $a < b$. Tällöin joukon A^* puujärjestys alkaa merkkijonoilla $\varepsilon, a, b, aa, ab, ba, bb, \ldots$.



Kuva 2.1: Esimerkin 2.4 ensimmäiset kaksi tasoa.

Osoitetaan, että *supistaminen* on voimassa joukossa A^* ja todistetaan Levin lemma merkkijonoille, jota hyödynnetään tutkittaessa merkkijonojen vaihdannaisuutta.

Lause 2.1. *Jos $x, y, z \in A^*$, niin tällöin $xz = yz$ implikoi $x = y$ ja vastaavasti $zx = zy$ implikoi $x = y$.*

Todistus. Oletetaan, että $x, y, z \in A^*$ ja olkoot $x = a_1a_2 \ldots a_l$, $y = b_1b_2 \ldots b_m$ ja $z = c_1c_2 \ldots c_n$. Jos nyt $xz = yz$, niin merkkijonot $a_1a_2 \ldots a_lc_1c_2 \ldots c_n$ ja $b_1b_2 \ldots b_mc_1c_2 \ldots c_n$ ovat samat. Koska merkkijono c on molempien merkkijonojen jälkiosa, niin merkkijonot $a_1a_2 \ldots a_l$ ja $b_1b_2 \ldots b_m$ ovat myös samoja, joten $x = y$. □

Lause 2.2 (Levin lemma)**.** *Olkoot $x, y, u, v \in A^*$. Oletetaan, että $xy = uv$. Jos $|x| \geq |u|$, niin on olemassa sellainen merkkijono $w \in A^*$, että $x = uw$ ja $v = wy$. Jos $|x| \leq |u|$, niin on olemassa sellainen merkkijono $w \in A^*$, että $u = xw$ ja $y = wv$.*

Todistus. Oletetaan, että $xy = uv = a_1a_2 \ldots a_l$. Tällöin merkkijonot x ja z ovat merkkijonon $a_1a_2 \ldots a_l$ etuosia eli $x = a_1a_2 \ldots a_i$ ja $u = a_1a_2 \ldots a_j$, missä $0 \leq i \leq l$ ja $0 \leq j \leq l$. Jos $i \geq j$, niin $x = uw$, missä $w = a_{j+1} \ldots a_i$, ja $v = a_{j+1} \ldots a_ia_{i+1} \ldots a_l = wy$. Jos taas $i \leq j$, niin $u = xw$, missä $w = a_{i+1} \ldots a_j$, ja $y = a_{i+1} \ldots a_l = a_{i+1} \ldots a_ja_{j+1} \ldots a_l = wv$. □

Seuraus 2.1. *Olkoot $x, y, u, v \in A^*$ ja $xy = uv$. Jos $|x| = |u|$, niin $x = u$ ja $y = v$.*

On selvää, että jos vähintään toinen merkkijonoista u tai v on tyhjä, niin $uv = vu$ ja tarkastellaan sen vuoksi rajoitettua tapausta, jossa $u, v \in A^+$.

Lause 2.3. *Olkoot $u, v \in A^+$. Tällöin $uv = vu$, jos ja vain jos on olemassa sellainen $z \in A^+$, että $u = z^p$ ja $v = z^q$ joillakin $p, q > 0$.*

Todistus. Suunta oikealta vasemmalle pätee selvästi.

Oletetaan sitten, että $uv = vu$. Jos $|u| = |v|$, niin seurauksen 2.1 nojalla $u = v$ ja väite pätee. Oletetaan symmetrian vuoksi, että $|u| > |v|$, jolloin $u \neq v$. Lauseen 2.2 perusteella on olemassa sellainen $z \in A^+$, että $u = vz = zv$ ja edelleen $u, v \in z^*$, joten väite pätee. $\square$

2.2 Kielet

Määritelmä 2.9. Olkoon A mikä tahansa aakkosto. Tällöin mitä tahansa aakkoston A kaikkien merkkijonojen joukon A^* osajoukkoa kutsutaan *kieleksi.*

Esimerkki 2.5. Koulumatematiikassa pärjätään aakkostolla $A = \{0, \ldots, 9\} \cup \{+, \cdot, -, /, =\} \cup \{(,)\}$, josta voidaan muodostaa kieli L vaikkapa kaikkien tosien kertolaskujen suhteen. Esimerkiksi $2 \cdot 2 = 4$ kuuluu tällöin kieleen L, mutta $0 \cdot 1 = 1$ ei kuulu kieleen L.

Määritelmä 2.10. Olkoon A aakkosto. Kieli on mikä tahansa joukon A^* osajoukko, joten $\mathcal{P}(A^*)$ *on kaikkien aakkostosta A muodostettujen kielten joukko.* Jos L ja M ovat aakkoston A kieliä, niin ovat myös $L \cup M$, $L \cap M$ ja $L \setminus M$. Jos L on aakkoston A kieli, niin on myös $L' = A^* \setminus L$. Merkitään yleisesti $L + M$ merkinnän $L \cup M$ sijaan. Näin ollen, jos L_i on *kieliperhe*, missä $1 \leq i \leq n$, niin niiden yhdistettä merkitään summana $\sum_{i=1}^{n} L_i$.

Määritelmä 2.11. Olkoon A aakkosto. Jos L ja M ovat sen kieliä, niin näiden *tulo* on $L \cdot M = \{xy \mid x \in L, y \in M\}$. Yleisesti merkitään LM merkinnän $L \cdot M$ sijaan.

Esimerkki 2.6. Huomataan aluksi, että $\emptyset L = \emptyset = L\emptyset$ ja $\{\varepsilon\}L = L = L\{\varepsilon\}$ pätevät millä tahansa kielellä L. Olkoot $L = \{ab, ba\}$ ja $M = \{aa, ab, ba, bb\}$. Nyt

$$LM = \{abaa, abab, abba, abbb, baaa, baab, baba, bb\}$$

ja

$$ML = \{aaab, aaba, abab, abba, baab, baba, bbab, bbab\},$$

joten $LM \neq ML$. Kielien tulo ei siis ole yleisesti vaihdannainen.

Määritelmä 2.12. Olkoon A aakkosto. Kielelle L määritellään $L^0 = \{\varepsilon\}$ ja $L^{n+1} = L^n \cdot L$. Jokaisella $n > 0$ kieli L^n muodostuu kaikista muotoa $u = x_1 x_2 \ldots x_n$ olevista merkkijonoista, missä $x_i \in L$.

Määritelmä 2.13. Olkoon A aakkosto. Kielen L *Kleenen tähti* on operaattori $*$, joka määritellään asettamalla $L^* = L^0 + L^1 + L^2 + \cdots + L^n$. Lisäksi määritellään operaattori $+$ asettamalla $L^+ = L^1 + L^2 + \cdots + L^n$.

Esimerkki 2.7. Sovitaan, että $\emptyset^* = \{\varepsilon\}$ ja $\{\varepsilon\}^* = \{\varepsilon\}$. Olkoon $A_1 = \{a\}$ aakkosto ja $L_1 = \{a\}^*$ kieli. Tällöin $L_1 = \{\varepsilon, a, a^2, a^3, \ldots\}$. Voidaan myös merkitä a^*. Olkoon sitten $A_2 = \{a, b\}$ aakkosto ja $L_2 = \{aa, bb\}^*$ kieli. Nyt $L_2 = \{\varepsilon, aa, bb, aaaa, bbbb, \ldots\}$, joten esimerkiksi $abab$ ei kuulu kieleen L_2. Kaikki kielen L_2 merkkijonot lukuunottamatta tyhjää merkkijonoa ε voidaan jakaa tekijöihin aa ja bb.

Määritelmä 2.14. Olkoon $A = \{a, b\}$. Yksinkertaistetaan kielien merkintää jättämällä hakasulkeet pois ja käytetään kaarisulkeita vain tarvittaessa. Tällöin pätee $A^* = \{a, b\}^* = (\{a\} + \{b\})^* = (a + b)^*$.

Esimerkki 2.8. Olkoon $A = \{a, b\}$ aakkosto. Kieli $L_1 = (a + b)^2$ koostuu kaikista neljästä kahden mittaisesta merkkijonosta. Kieli $L_2 = ab(a + b)^*$ koostuu kaikista merkkijonoista, jotka alkavat merkkijonolla ab. Kieli

$$L_3 = (a + b)^* a (a + b)^* b (a + b)^* b (a + b)^* a (a + b)^*$$

koostuu niistä merkkijonoista, jotka sisältävät osamerkkijonona merkkijonon $abba$. Kieli $L_4 = (a + b)^* baba (a + b)^*$ koostuu kaikista niistä merkkijonoista, jotka sisältävät tekijänä merkkijonon $baba$. Kieli $L_5 = (aa + ab + ba + bb)^*$ koostuu kaikista niistä merkkijonoista, joiden pituus on parillinen.

Lause 2.4. *Olkoon L mikä tahansa kieli. Jos $x, y \in L^*$, niin $xy \in L^*$.*

Todistus. Oletetaan, että $xy \notin L^*$. Koska L^* on yhdiste, niin $xy \notin L^i$ jokaisella $i \in \mathbb{N}_0$. Siis $x \notin L^i$ tai $y \notin L^i$ jokaisella $i \in \mathbb{N}_0$, joten $x, y \notin L^*$. $\square$

Lause 2.5. *Olkoot* $L, M, N \subseteq A^*$*. Tällöin seuraavat ominaisuudet pätevät kielillä* L, M *ja* N*:*

i) $L(MN) = (LM)N$

ii) $L(M + N) = LM + LN$ *ja* $(M + N)L = ML + NL$

iii) Jos $L \subseteq M$*, niin* $NL \subseteq NM$ *ja* $LN \subseteq MN$*.*

Todistus. Tapauksien todistuksissa sovelletaan lähes pelkästään määritelmää 2.11 ja yhdisteen määritelmää, joten ne sivuutetaan ilmeisinä. □

2.3 Deterministinen äärellinen automaatti

Määritelmä 2.15. *Deterministinen äärellinen automaatti* on viisikko $\mathbf{A} = (S, A, i, \delta, T)$, missä S on äärellinen *tilojen* joukko, A on aakkosto, $i \in S$ on *alkutila*, δ on *siirtymäkuvaus* $\delta : S \times A \to S$ ja $T \subseteq S$ on *lopputilojen* joukko. *Siirtymä* on tapahtuma, jossa automaatti vaihtaa jollakin joukon A *syötteellä* a tilan jostakin tilasta s johonkin tilaan t eli $\delta(s, a) = t$. Kutsutaan determinististä automaattia lyhyemmin automaatiksi ja merkitään $\mathbf{A}^{det}$. Lisäksi merkinnän $\delta(s, a)$ sijaan käytetään yleisesti merkintää sa.

Määritelmä 2.16. Olkoon $\mathbf{A}^{det}$ automaatti. Määritellään seuraavaksi *laajennettu siirtymäkuvaus* δ^*, jonka avulla $\mathbf{A}^{det}$ voi käsitellä myös merkkijonoja. $\delta^* : S \times A^* \to S$ on yksikäsitteinen kuvaus, joka toteuttaa seuraavat ehdot, kun $a \in A$, $x \in A^*$ ja $s \in S$:

i) $\delta^*(s, \varepsilon) = s$

ii) $\delta^*(s, a) = \delta(s, a)$

iii) $\delta^*(s, ax) = \delta^*(\delta(s, a), x)$.

Automaatti $\mathbf{A}^{det}$ on deterministinen, koska nykyinen tila ja merkki yksikäsitteisesti määräävät uuden tilan. Lisäksi automaattia $\mathbf{A}^{det}$ kutsutaan *täydelliseksi*, koska jokaisesta tilasta on siirtymä johonkin tilaan.

Lause 2.6. *Olkoon S äärellinen joukko ja A äärellinen aakkosto. Olkoon $\delta : S \times A \to S$ kuvaus. Tällöin on olemassa yksikäsitteinen kuvaus $\Delta : S \times A^* \to S$, joka toteuttaa seuraavat kolme ehtoa:*

i) $\Delta(s, \varepsilon) = s$, *kun* $s \in S$

ii) $\Delta(s, a) = \delta(s, a)$, *kun* $s \in S$ *ja* $a \in A$

iii) $\Delta(s, ax) = \Delta(\delta(s, a), x)$, *kun* $s \in S$, $a \in A$ *ja* $x \in A^*$.

Todistus. Kts. [3, ss. 17-18]. □

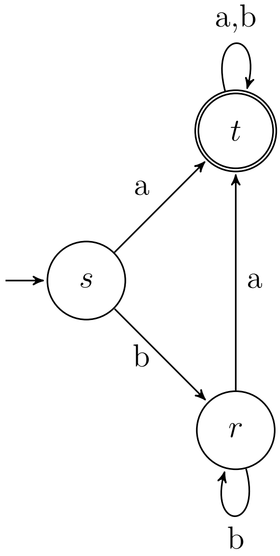
Lause 2.7. *Olkoon* $\mathbf{A}^{det}$ *automaatti. Tällöin* $s(xy) = (sx)y$ *jokaisella* $x, y \in A^*$ *ja* $s \in S$.

Todistus. Kts. [3, s. 18]. □

Määritelmä 2.17. Olkoon $\mathbf{A}^{det}$ automaatti ja olkoot $a \in A, s, t \in S$. *Siirtymäverkko* on suunnattu verkko, jonka solmut nimetään automaatin $\mathbf{A}^{det}$ tilojen mukaan ja nuolella $s \to t$ on *nimike* a ainoastaan silloin, kun $\delta(s, a) = t$. Lisäksi alkutilaan merkitään sisäänpäin osoittava nimikkeetön nuoli ja lopputilassa on kaksinkertainen ympyrä. Sovitaan, että mikäli merkit $a_1, a_2, \ldots, a_n$ nimittävät n siirtymää tilasta s tilaan t, niin riittää piirtää vain yksi nuoli $s \to t$, jonka nimike on $a_1, a_2, \ldots, a_n$.

Määritelmä 2.18. Olkoon $\mathbf{A}^{det}$ automaatti. *Siirtymätaulukko* on taulukko, jossa kuvauksen δ kuvat esitetään siten, että riveillä ovat tilat ja sarakkeissa ovat merkit. Alkutilaa merkitään nuolella vasemmalta oikealle ja lopputilaa nuolella oikealta vasemmalle.

Esimerkki 2.9. $\mathbf{A}^{det} = (\{s, t, r\}, \{a, b\}, s, \delta, \{t\})$ on automaatti, jonka siirtymät ovat $sa = t$, $sb = r$, $ra = t$, $rb = r$, $ta = t$ ja $tb = t$. Automaatin $\mathbf{A}^{det}$ siirtymäverkko on kuvassa 2.2 ja siirtymätaulukko taulukossa 2.1.



Kuva 2.2: Esimerkin 2.9 siirtymäverkko.

	a	b
$\rightarrow s$	t	r
$\leftarrow t$	t	t
r	t	r

Taulukko 2.1: Esimerkin 2.9 siirtymätaulukko.

Määritelmä 2.19. Automaatti $\mathbf{A}^{det}$ on *hyvin muodostettu*, mikäli sillä on äärellinen määrä tiloja ja se toteuttaa seuraavat ehdot:

i) Automaatilla on vain yksi alkutila.

ii) Jokaisella tilalla s ja jokaisella syötteellä $a \in A$ on olemassa vain yksi nuoli tilasta s tilaan sa ja jonka nimike on a.

Määritelmä 2.20. Olkoon $\mathbf{A}^{det} = (S, A, i, \delta, T)$ automaatti. Määritellään automaatin $\mathbf{A}^{det}$ *tunnistama kieli* $L(\mathbf{A}^{det})$ asettamalla $L(\mathbf{A}^{det}) = \{x \in A^* \mid ix \in T\}$. Kielen sanotaan olevan *tunnistettavissa*, mikäli on olemassa automaatti, joka tunnistaa sen.

Automaatin $\mathbf{A}^{det}$ tunnistama kieli muodostuu siis kaikista niistä merkkijonoista, joiden merkeillä voidaan merkitä polku alkutilasta lopputilaan. Tällöin sanotaan, että automaatti $\mathbf{A}^{det}$ *hyväksyy* merkkijonon. Tyhjän merkkijonon tapauksessa tarvitaan ehto, joka käy ilmi seuraavasta tuloksesta.

Lause 2.8. *Olkoon* $\mathbf{A}^{det} = (S, A, i, \delta, T)$ *automaatti ja* $L(\mathbf{A}^{det})$ *sen tunnistama kieli. Tällöin automaatti* $\mathbf{A}^{det}$ *hyväksyy tyhjän merkkijonon, jos ja vain jos* $i \in T$.

Todistus. Oletetaan, että $i \notin T$. Koska määritelmän 2.16 nojalla $i = i\varepsilon$, niin määritelmän 2.20 nojalla $\varepsilon \notin L(\mathbf{A}^{det})$.

Oletetaan sitten, että $\varepsilon \notin L(\mathbf{A}^{det})$. Täten määritelmän 2.20 nojalla $i\varepsilon \notin T$, joten $i \notin T$. □

Esimerkki 2.10. Monessa tapauksessa automaatin hyväksymän kielen näkee helposti siirtymäverkosta, mutta sen osoittaminen formaalisti on työlästä. Osoitetaan, että esimerkin 2.9 automaatti hyväksyy kielen $(a + bb^*)(a + b)^*$.

Aluksi havaitaan, että mikä tahansa merkkijono $x \in A^*$, joka alkaa merkillä a, hyväksytään tilassa t. Huomataan myös, että jotkut merkkijonot $y \in A^*$, jotka alkavat merkillä b, hyväksytään tilassa t.

Olkoon $x \in A^*$ sellainen merkkijono, että $x = ax'$. Jos ollaan tilassa s, niin syötteellä a päädytään tilaan t, jossa myös pysytään oli merkkijonon loppuosa mitä tahansa.

Olkoon sitten $y \in A^*$ sellainen merkkijono, että $y = by'$. Jos ollaan tilassa s, niin syötteellä b päädytään tilaan r. Mikäli merkkijonon y' ensimmäinen merkki on a, niin päädytään tilaan t ja pysytään siellä oli merkkijonon loppuosa mitä tahansa. Mikäli merkkijonon y' ensimmäinen merkki on b, niin pysytään tilassa r. Tilassa r pysytään, kunnes luetaan a ja päädytään tilaan t, jossa pysytään oli merkkijonon loppuosa mitä tahansa.

Lopuksi lauseen 2.8 nojalla tyhjää merkkijonoa ei hyväksytä, koska alkutila s ei kuulu lopputilojen joukkoon. Näin on osoitettu, että $L(\mathbf{A}^{det}) = (a + bb^*a)(a + b)^*$.

Kun sanotaan, että automaatti $\mathbf{A}^{det}$ tunnistaa kielen L, niin tällöin $L = L(\mathbf{A}^{det})$. Lawson [3, luku 2] esittelee lukuisia esimerkkejä erilaisista tunnistettavista kielistä. Toisaalta voidaan osoittaa, että kieli $L = \{a^n b^n \mid n \in \mathbb{N}\}$ ei ole tunnistettavissa [3, s. 35]. Seuraavaa tulosta voidaan käyttää tutkittaessa, onko jokin kieli L tunnistamaton.

Lause 2.9 (Pumppauslemma)**.** *Olkoon $L \subseteq A^*$ tunnistettava kieli. Tällöin on olemassa sellainen $n \in \mathbb{N}$, että jokaisella sellaisella $w \in L$, että $|w| \geq n$, on olemassa sellaiset merkkijonot $x, y, z \in L$, että seuraavat ehdot pätevät:*

i) $w = xyz$

ii) $|xy| \leq n$

iii) $|y| \geq 1$

iv) Jokaisella $i \geq 0$ pätee $xy^iz \in L$.

Todistus. Kts. [3, s. 46]. □

Esimerkki 2.11. Osoitetaan pumppauslemman 2.9 avulla, että kieli $L = \{a^n b^{2n} \mid n \geq 0\}$ ei ole tunnistettavissa. Tehdään vastaoletus, että L on tunnistettavissa. Tällöin on olemassa sellainen $n \in \mathbb{N}$, että jokaisella sellaisella $w \in L$, että $|w| \geq n$, on olemassa sellaiset merkkijonot $x, y, z \in L$, että pumppauslemman ehdot pätevät: (i) $a^n b^{2n} = xyz$, (ii) $|xy| \leq n$, (iii) $|y| \geq 1$ ja (iv) $xy^iz \in L$ jokaisella $i \geq 0$.

Kohdan (ii) nojalla merkkijono xy koostuu pelkistä a merkeistä ja kohta (iii) kertoo, että merkkijonossa y on ainakin yksi a merkki. Siispä merkkijonossa z on kaikki b-merkit ja mahdollisesti a-merkkejä.

Tarkastellaan merkkijonoa $xz = xy^0z$, joka kuuluu kieleen L pumppauslemman nojalla. Huomataan, että $|xz|_b = 2n$, mutta $|xz|_a = n - |y|_a < n$, joten $xz \notin L$. Päädyttiin ristiriitaan, joten vastaoletus on väärin, eikä L ole tunnistettavissa.

Määritelmä 2.21. Olkoon $\mathbf{A}^{det} = (S, A, i, \delta, T)$ automaatti. Sanotaan, että tila $s \in S$ on *saavutettava*, mikäli on olemassa sellainen merkkijono $x \in A^*$, että $ix = s$. Tilaa, joka ei ole saavutettavissa, kutsutaan *saavuttamattomaksi.* Automaatti $\mathbf{A}^{det}$ on *saavutettava*, mikäli kaikki sen tilat ovat saavutettavia automatin alkutilasta.

Alkutila on aina saavutettavissa, koska $i\varepsilon = i$. Huomataan myös, että saavuttamattomilla tiloilla ei ole merkitystä automaatin toiminnan kannalta.

Määritelmä 2.22. Olkoon $\mathbf{A}^{det} = (S, A, i, \delta, T)$ automaatti. Määritellään sen avulla *saavutettava automaatti* $\mathbf{A}^a = (S^a, A^a, i^a, \delta^a, T^a)$, missä $S^a = S$ on tilojen joukko, $i^a = i$ on alkutila, $T^a = T \cap S^a$ on lopputilojen joukko ja $\delta^a : S^a \times A \to S^a$ on siirtymäkuvaus, joka käyttäytyy kuten δ. Usein puhutaan lyhyemmin automaatista $\mathbf{A}^a$.

Lause 2.10. *Olkoon $\mathbf{A}^{det} = (S, A, i, \delta, T)$ automaatti. Tällöin $\mathbf{A}^a$ on saavutettava automaatti ja $L(\mathbf{A}^{det}) = L(\mathbf{A}^a)$.*

Todistus. Kts. [3, s. 54]. □

Saavutettavaa automaattia $\mathbf{A}^a$ sanotaan automaatin $\mathbf{A}^{det}$ *saavutettavaksi osaksi.* On yksinkertaista muodostaa saavutettava automaatti, kun tilojen lukumäärä on pieni. Tilojen lukumäärän kasvaessa suureksi tilanne muuttuu hankalammaksi.

Lause 2.11. *Olkoon $\mathbf{A}^{det} = (S, A, s_0, \delta, T)$ automaatti. Jos $s \in S$ on hyväksyttävä tila, niin on olemassa sellainen merkkijono $x \in A^*$, että $|x| < |S|$ ja $s_0 x = s$.*

Todistus. Kts. [3, ss. 55-56]. □

Lauseen 2.11 nojalla saavutettavan automaatin $\mathbf{A}^a$ tilojen joukko S^a voidaan selvittää seuraavalla tavalla. Olkoon $|S| = n$ ja s_0 alkutila. Olkoot lisäksi $X \subseteq S$ ja $L \subseteq A^*$ ja määritellään $XL = \{xa \mid x \in X, a \in L\}$. Kaikkien merkkijonojen $x \in A^*$ joukkoa, joille pätee $|x| \leq n - 1$, merkitään $\sum_{i=0}^{n-1} A^i$. Lauseen 2.11 tulos voidaan siis ilmaista muodossa

$$S^a = s_0 \left(\sum_{i=0}^{n-1} A^i \right) = \sum_{i=0}^{n-1} s_0 A^i.$$

Joukkojen $S_i = s_0 A^i$ avulla voidaan muodostaa yllä oleva yhdiste. Tätä voidaan tarkastella puista koostuvan jonon avulla. Kahden solmun välisellä etäisyydellä tarkoitetaan lyhintä polkua näiden kahden solmun välillä. Puun korkeus on pisimmän sellaisen merkkijonon pituus juuresta lehteen, joka ei käy samassa solmussa useammin kuin kerran. Puun juuri on alkutila s_0. Jokaisella $a \in A$ muodostetaan nuoli $s_0 \to s_0 a$.

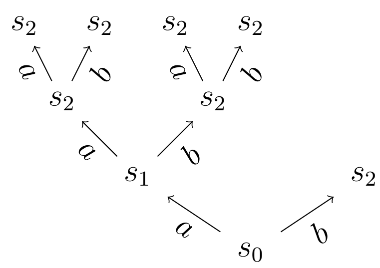
Jos s on solmu, niin piirretään nuolia $s \to sa$, joiden nimike on a, jokaisella $a \in A$. Solmut, jotka ovat etäisyyden i päässä juuresta, ovat joukon $s_0 A^i$

alkioita. Lopetetaan, kun puun korkeudeksi on saatu $n-1$. Puun solmut ovat nyt ne automaatin tilat, jotka ovat saavutettavia.

Esimerkki 2.12. Olkoon $\mathbf{A}^{det} = (\{s_0, s_1, s_2, s_3\}, \{a, b\}, s_0, \delta, \{s_1\})$ automaatti, jonka siirtymät ovat taulukossa 2.2. Koska $|\{s_0, s_1, s_2, s_3\}| = 4$, muodostetaan puu, jonka korkeus on kolme ja esitetään se kuvassa 2.3.

	a	b
$\rightarrow s_0$	s_1	s_2
$\leftarrow s_1$	s_2	s_2
s_2	s_2	s_2
s_3	s_1	s_1

Taulukko 2.2: Esimerkin 2.12 siirtymätaulukko.



Kuva 2.3: Esimerkin 2.12 siirtymäpuu.

Automaatti $\mathbf{A}^a$ voidaan nyt muodostaa automaatista $\mathbf{A}^{det}$ poistamalla kaikki saavuttamattomat solmut ja nuolet, jotka menevät niihin tai lähtevät niistä. Jos automaatissa on n tilaa, niin täytyy piirtää puu, jonka korkeus on $n-1$. Tämä on työlästä ja koska solmu voi esiintyä puussa useammin kuin kerran, niin puussa voi olla toistoa, joka ei tuota uutta informaatiota. Seuraava algoritmi hylkää joukon sA, jos s toistuu.

Määritelmä 2.23. Olkoon $\mathbf{A}^{det} = (S, A, i, \delta, T)$ automaatti. Muodostetaan automaatin $\mathbf{A}^{det}$ *siirtymäpuu* induktiivisesti seuraavalla tavalla. Jos sa on minkä tahansa solmun toisto, niin sanotaan, että *solmu on suljettu* ja merkitään $\times$. Lisäksi tässä algoritmissa ja seuraavassa lauseessa siirtymäpuuta merkitään T_i, missä $i \geq 0$. Oletetaan, että aakkostossa A pätee lineaarijärjestys. Algoritmi voidaan esittää seuraavasti:

i) Puun juuri on s_0, ja asetetaan $T_0 = \{s_0\}$.

ii) Oletetaan, että T_i on muodostettu eli kaikki sen solmut on merkitty joko suljetuiksi tai ei-suljetuiksi.

iii) Jokaisella ei-suljetulla solmulla $s \in T_i$ ja jokaisella $a \in A$ lisätään nuoli $s \to sa$, jonka merkki on a. Sovitaan, että jos $sa = sb = p$, niin riittää yksi nuoli $s \to p$, jonka nimike on a, b.

iv) Algoritmi loppuu, kun kaikki lehdet ovat suljettuja.

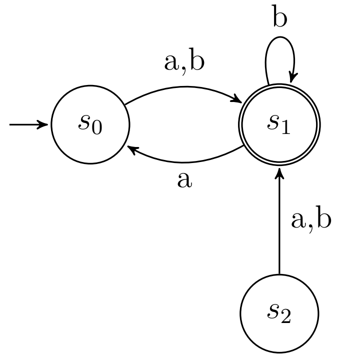
Seuraava lause osoittaa, että määritelmän 2.23 algoritmi toimii oikein.

Lause 2.12. *Olkoon* $|S| = n$. *Tällöin on olemassa sellainen* $m \leq n$, *että jokainen puun* T_m *solmuista on suljettu ja* S^a *on vain puun* T_m *solmujen joukko.*

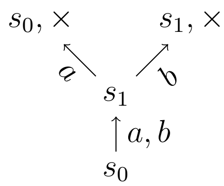
Todistus. Kts. [3, s. 57]. Olkoot $s \in S^a$ ja $x \in A^*$ pienin sellainen merkkijono puujärjestyksessä, että $s_0 x = s$. Tällöin s esiintyy ensimmäisenä puun $T_{|x|}$ solmuna. Lauseen 2.11 nojalla saavutettava tila voi esiintyä ensimmäisen kerran viimeistään puussa T_{n-1}. Siten kaikki T_n solmut ovat suljettuja. □

Siirtymäpuun avulla voidaan selvittää, mitkä ovat automaatin $\mathbf{A}^{det}$ saavutettavat tilat, mutta sen avulla voidaan myös muodostaa saavutettava automaatti $\mathbf{A}^a$. Tämä onnistuu poistamalla kaikki merkit $\times$ ja liimaamalla lehdet edeltäviin solmuihin, joilla on sama merkki.

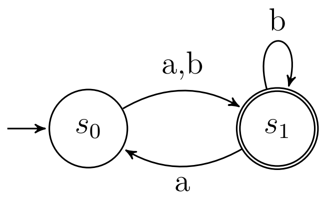
Esimerkki 2.13. Olkoon $\mathbf{A}^{det} = (S, A, s_0, \delta, T)$ automaatti, missä $S = \{s_0, s_1, s_2\}$, $A = \{a, b\}$, $T = \{s_1\}$, ja siirtymät ovat $s_0 a = s_1$, $s_0 b = s_1$, $s_1 a = s_0$, $s_1 b = s_1$, $s_2 a = s_1$ ja $s_2 b = s_1$. Automaatin $\mathbf{A}^{det}$ siirtymäverkko on esitetty kuvassa 2.4. Sivuutetaan siirtymäpuut T_0 ja T_1, ja piirretään vain T_2, joka on esitetty kuvassa 2.5. Lopuksi saadaan automaatti $\mathbf{A}^a = (S^a, A^a, i^a, \delta^a, T^a)$, joka on muodostettu määritelmän 2.22 vaatimalla tavalla ja esitetty kuvassa 2.6.



Kuva 2.4: Esimerkin 2.13 siirtymäverkko.



Kuva 2.5: Esimerkin 2.13 siirtymäpuu T_2.



Kuva 2.6: Esimerkin 2.13 siirtymäverkko.

2.4 Epädeterministinen äärellinen automaatti

Tarkastellaan seuraavaksi automaattia, jonka siirtymät ovat epädeterministisiä. Epädeterministisyys viittaa siihen, että siirtymät eivät ole deterministisiä eli samalla merkillä voi päästä siirtyä tilasta useaan tilaan. Lisäksi on muutakin, joka eroaa deterministisen äärellisen automaatin määritelmästä: alkutiloja voi olla enemmän kuin yksi.

Määritelmä 2.24. *Epädeterministinen automaatti* on $\mathbf{A}^{edet} = (S, A, I, \delta, T)$, missä S on äärellinen *tilojen* joukko, A on aakkosto, I on *alkutilojen* joukko,

$\delta : S \times A \to \mathcal{P}(S)$ on *siirtymäkuvaus*, ja T on *lopputilojen* joukko. Usein puhutaan lyhyemmin automaatista $\mathbf{A}^{edet}$.

Epädeterministisen automaatin $\mathbf{A}^{edet}$ siirtymäverkko- ja taulukko muodostetaan kuten automaatin $\mathbf{A}^{det}$ tapauksessa. Esitetään määritelmästä 2.16 uusi versio vastaamaan epädeterministisen automaatin $\mathbf{A}^{edet}$ siirtymäkuvausta δ.

Määritelmä 2.25. Olkoon $\mathbf{A}^{edet}$ epädeterministinen automaatti. *Epädeterministisen automaatin laajennettu siirtymäkuvaus* on yksikäsitteinen kuvaus $\delta^* : S \times A^* \to \mathcal{P}(S)$, joka toteuttaa seuraavat ehdot, kun $a \in A$, $x \in A^*$ ja $s \in S$:

i) $\delta^*(s, \varepsilon) = \{s\}$

ii) $\delta^*(s, a) = \delta(s, a)$

iii) $\delta^*(s, ax) = \sum_{q \in \delta(s,a)} \delta^*(q, x)$.

Huomautus. Merkitään sx merkinnän $\delta^*(s, x)$ sijaan, mutta täytyy muistaa, että sx on joukko, kun on kyse epädeterministisestä automaatista. Lisäksi käytetään merkitää $\{s\}$ merkinnän s sijasta.

Lause 2.13. *Olkoon $x = a_1 a_2 \cdots a_n \in A^*$. Tällöin $t \in sx$, jos ja vain jos on olemassa sellaiset tilat $q_1, \ldots, q_n = t$, että $q_1 \in sa_1$, ja $q_i \in q_{i-1} a_i$, jokaisella $2 \leq i \leq n$.*

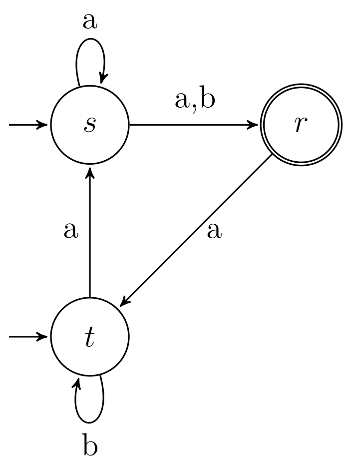
Todistus. Kts. [3, s. 62]. □

Määritelmä 2.26. Automaatin $\mathbf{A}^{edet}$ *tunnistama kieli* $L(\mathbf{A}^{edet})$ on

$$L(\mathbf{A}^{edet}) = \{x \in A^* \mid \left(\sum_{q \in I} qx \right) \cap T \neq \emptyset\}.$$

Esimerkki 2.14. $\mathbf{A}^{edet} = (\{s, r, t\}, \{a, b\}, \{s, t\}, \delta, r)$ on epädeterministinen automaatti, jonka siirtymät ovat $sa = \{s, r\}$, $sb = r$, $ta = s$, $tb = t$ ja $ra = t$. Sen siirtymätaulukko on esitetty taulukossa 2.3 ja siirtymäverkko kuvassa 2.7.

Epädeterminismin voisi luulla tuovan lisää ilmaisuvoimaa, mutta äärellisten automaattien tapauksessa näin ei ole. Olkoon $\mathbf{A}^{edet} = (S, A, I, \delta, T)$ epädeterministinen automaatti. Muodostetaan deterministinen automaatti $\mathbf{A}^{d} =$



Kuva 2.7: Esimerkin 2.14 siirtymäverkko.

	a	b
$\rightarrow s$	$\{s, r\}$	r
$\rightarrow t$	s	t
$\leftarrow r$	t	

Taulukko 2.3: Esimerkin 2.14 siirtymätaulukko.

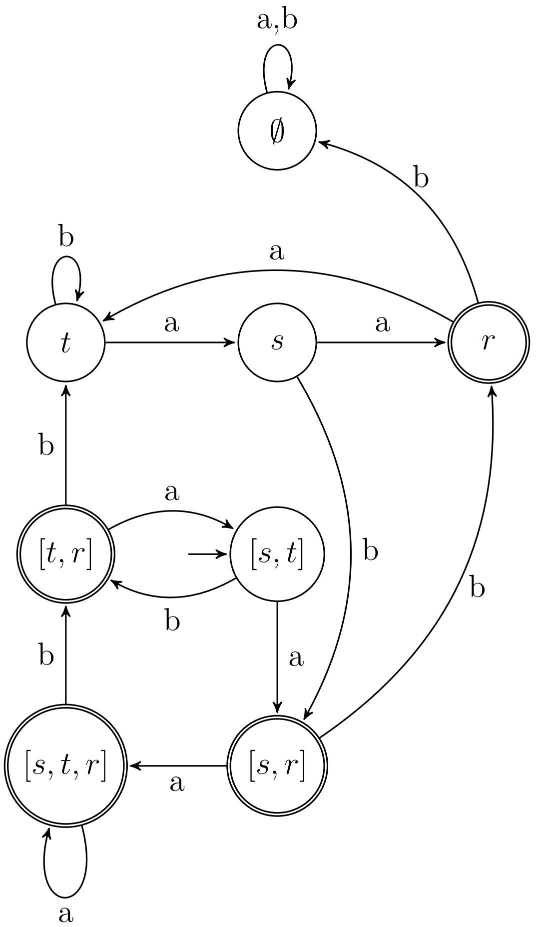
$(S^d, A^d, i^d, \Delta, T^d)$, missä $S^d = \mathcal{P}(S)$, alkutila $i^d = I$ on epädeterministisen automaatin alkutilojen osajoukko $I \subseteq \mathcal{P}(S)$, $T^d = \{Q \in \mathcal{P}(S) \mid Q \cap T \neq \emptyset\}$. Lisäksi jokaisella $a \in A$ ja $Q \in \mathcal{P}(S)$ asetetaan $\Delta(Q, a) = \sum_{q \in Q} qa$. Joukko $\Delta(Q, a)$ koostuu kaikista niistä tiloista $s \in S$, jotka voidaan saavuttaa joukon Q tiloista merkillä a. Näin ollen $\mathbf{A}^d$ on täydellinen deterministinen automaatti $\mathbf{A}^d$.

Lause 2.14. *Olkoon $\mathbf{A}^{edet}$ epädeterministinen automaatti. Tällöin on olemassa sellainen deterministinen automaatti $\mathbf{A}^{det}$, että $L(\mathbf{A}^d) = L(\mathbf{A}^{edet})$.*

Todistus. Kts. [3, ss. 63-64]. Todistuksessa osoitetaan induktiolla $|x|$ suhteen, että jokaisella $Q \subseteq S$ ja $x \in A^*$ pätee $\Delta^*(Q, x) = \sum_{q \in Q} \delta^*(q, x)$. Lisäksi $L(\mathbf{A}^d) = L(\mathbf{A}^{edet})$ seuraa suoraan määritelmistä ja induktiolla osoitetusta indentiteetistä. $\square$

Automaattia $\mathbf{A}^d$ kutsutaan automaatin $\mathbf{A}^{edet}$ *determinoiduksi automaatiksi.* Automaatin $\mathbf{A}^d$ muodostamista kutsutaan *determinisoinniksi.* Myös tässä tapauksessa tilojen lukumäärän kasvu johtaa hyvin nopeasti vaikeuksiin. Jos automaatissa $\mathbf{A}^{edet}$ on n tilaa, niin automaatissa $\mathbf{A}^{d}$ on 2^n tilaa. Lawson [3, s. 65] antaa algoritmin, jonka avulla determinisointi onnistuu helposti.

Esimerkki 2.15. Esitetään esimerkin 2.14 automaatista determinisoitu versio kuvassa 2.8.



Kuva 2.8: Esimerkin 2.14 automaatin $\mathbf{A}^{edet}$ determinisoidun automaatin $\mathbf{A}^{d}$ siirtymäverkko.

Lawson [3, ss. 67-70] esittelee useita epädeterministisellä äärellisellä automaatilla tunnistettavia kieliä.

2.5 Säännölliset kielet

Määritelmä 2.27. Olkoon $A = \{a_1, a_2, \ldots, a_n\}$ aakkosto. Määritellään, että *säännöllinen lauseke* aakkostossa A on merkkijono, joka muodostetaan soveltamalla äärellinen määrä kertoja seuraavia sääntöjä:

i) $\emptyset$ on säännöllinen lauseke

ii) ε on säännöllinen lauseke

iii) $a_1, a_2, \ldots, a_n$ ovat kaikki säännöllisiä lausekkeita

iv) Jos s ja t ovat säännöllisiä lausekkeita, niin on myös $(s + t)$

v) Jos s ja t ovat säännöllisiä lausekkeita, niin on myös $(s \cdot t)$

vi) Jos s on säännöllinen lauseke, niin on myös (s^*)

vii) Jokainen säännöllinen lauseke voidaan muodostaa äärellisellä lukumäärällä sääntöjä (i)-(vi).

Operaattoreita $+$, $\cdot$ ja $*$ kutsutaan *säännöllisiksi operaattoreiksi*. Yleensä $\cdot$ jätetään merkitsemättä. Sovitaan, että säännöllisten lausekkeiden sulkeita voidaan jättää pois kuten algebrallisten lausekkeiden tapauksessa siten, että $*$ vastaa potenssia, $\cdot$ vastaa kertolaskua ja $+$ vastaa yhteenlaskua. Toisin sanoen esimerkiksi $((0 \cdot (1^*)) + 0)$ voidaan lyhentää muotoon $01^* + 0$. Tällöin siis esimerkiksi 01^* vastaa merkintää $0(1^*)$, eikä merkintää $(01)^*$.

Määritelmä 2.28. Säännölliset lausekkeet s ja t ovat *samat*, merkitään $s = t$, jos ja vain jos $L(s) = L(t)$.

Määritelmä 2.29. Kieli L on *säännöllinen*, mikäli on olemassa sellainen säännöllinen lauseke s, että $L = L(s)$.

Seuraavien kolmen lauseen todistukset sivuutetaan, koska ne ovat suoraviivaisia relevanttien määritelmien soveltamisia.

Lause 2.15. *Jos L, M, N ovat kieliä, niin seuraavat identiteetit pätevät: $L(\emptyset) = \emptyset$, $L(\varepsilon) = \{\varepsilon\}$, $L(a_i) = \{a_i\}$, $L(s + t) = L(s) + L(t)$, $L(st) = L(s)L(t)$ ja $L(s^*) = L(s)^*$.*

Lause 2.16. *Mikäli r, s, t ovat säännöllisiä lausekkeita, niin pätee $r + (s + t) = (r + s) + t$, $r(st) = (rs)t$, $r(s + t) = rs + rt$, $r^* = (rr)^* + r(rr)^*$, $(r + s)^* = (r^*s^*)^*$ ja $(rs)^*r = r(sr)^*$.*

Lause 2.17. *Olkoon A aakkosto ja olkoot $L, M, N \in \mathcal{P}(A^*)$. Tällöin seuraavat ominaisuudet pätevät kielillä L, M ja N:*

i) $L + (M + N) = (L + M) + N$

ii) $\emptyset + L = L = L + \emptyset$

iii) $L + L = L$

iv) $L \cdot (M \cdot N) = (L \cdot M) \cdot N$

v) $\varepsilon \cdot L = L = L \cdot \varepsilon$

vi) $\emptyset \cdot L = \emptyset = L \cdot \emptyset$

vii) $L \cdot (M + N) = L \cdot M + L \cdot N$ *ja* $(M + N) \cdot L = M \cdot L + N \cdot L$.

Lauseen 2.17 ominaisuudet pätevät myös säännöllisille lausekkeille määritelmän 2.28 nojalla.

Esimerkki 2.16. Olkoon $L_1 = \{x \in (a + b)^* \mid |x| < 4\}$. Tällöin merkkijono $x \in L$, jos sen pituus on 0,1,2 tai 3, jolloin L voidaan ilmaista säännöllisellä lausekkeella $\varepsilon + (a + b) + (a + b)^2 + (a + b)^3$. Toisaalta taas säännöllisellä lausekkeella $(a+b)^4(a+b)^*$ voidaan ilmaista kieli $L_2 = \{x \in (a+b)^* \mid |x| \geq 4\}$.

2.6 Kleenen lause äärelliselle automaatille

Kleenen lause on merkittävä tulos äärellisten automaattien teoriassa ja todistetaan se tässä alaluvussa. Ensiksi kuitenkin esitellään todistuksessa tarvittavia tuloksia, joiden todistuksissa sovelletaan määritelmää 2.24. Näiden tuloksien todistukset löytyvät mm. Lawsonilta [3, s. 43, ss. 68-69 ja ss. 70-71].

Lause 2.18. *Jos L ja M ovat tunnistettavia kieliä, niin on myös $L + M$.*

Lause 2.19. *Jos L ja M ovat tunnistettavia kieliä, niin on myös LM.*

Lause 2.20. *Jos L on tunnistettava kieli, niin on myös L^*.*

Lause 2.21 (Kleenen lause)**.** *Kieli on tunnistettavissa, jos ja vain jos se on säännöllinen.*

Todistus. Vrt. [3, ss. 103-104] ja [2, ss. 58-63]. Näistä erityisesti jälkimmäinen on tarkempi, kun todistetaan väitettä vasemmalta oikealle.

Olkoon A aakkosto ja L siitä muodostettu säännöllinen kieli. Osoitetaan, että L on tunnistettavissa. Osoitetaan tämä induktiolla säännöllisessä lausekkeessa olevien säännöllisten operaattorien lukumäärän suhteen. Säännölliset lausekkeet, joissa ei ole säännöllisiä operaattoreita, voivat kuvata vain kielet ε, $\emptyset$ ja $\{a\}$, missä $a \in A$. Jokainen näistä kielistä on tunnistettavissa määritelmän 2.20 nojalla, koska voidaan helposti muodostaa automaatti, joka tunnistaa nämä kielet. Tämä on induktion perusaskel.

Induktio-oletus on, että jos r on säännöllinen lauseke, jossa on korkeintaan $n-1$ säännöllistä operaattoria, niin $L(r)$ on tunnistettavissa. Olkoon sitten r säännöllinen lauseke, jossa on n säännöllistä operaattoria. Tällöin on kolme vaihtoehtoa: $r = s + t$, $r = st$ ja $r = s^*$, missä s, t ovat säännöllisiä lausekkeita, joissa on korkeintaan $n-1$ säännöllistä operaattoria. Induktio-oletuksen nojalla $L(s)$ ja $L(t)$ ovat molemmat tunnistettavissa. Joten voidaan soveltaa lauseita 2.18, 2.19 ja 2.20, jolloin voidaan päätellä, että $L(r)$ on tunnistettavissa. Siis L on tunnistettavissa.

Oletetaan sitten, että L on tunnistettava kieli ja osoitetaan, että se on säännöllinen kieli. Käytetään epädeterminististä automaattia $\mathbf{A}^{edet}$, jonka siirtymäverkon nuolien kokonaismäärää kutsutaan automaatin $\mathbf{A}^{edet}$ *siirtymäluvuksi*. Osoitetaan väite induktiolla tämän luvun suhteen.

Jos automaatin $\mathbf{A}^{edet}$ siirtymäluku on nolla, niin tällöin $L(\mathbf{A}^{edet})$ on joko $\emptyset$ tai ε, joista jälkimmäinen toteutuu, jos jokin automaatin alkutiloista on myös lopputila. Tämä on induktion perusaskel.

Induktio-oletus on, että jos $\mathbf{A}^{edet}$ on epädeterministinen automaatti, jonka siirtymäluku on korkeintaan $n-1$, niin $L(\mathbf{A}^{edet})$ on säännöllinen.

Olkoon $\mathbf{A}^{edet} = (S, A, I, \delta, T)$ epädeterministinen automaatti, jonka siirtymäluku on n. Osoitetaan, että $L(\mathbf{A}^{edet})$ on säännöllinen. Oletuksen nojalla automaatin $\mathbf{A}^{edet}$ siirtymäverkossa on vähintään yksi siirtymä. Valitaan yksi mielivaltainen siirtymä $p \xrightarrow{a} q$. Muodostetaan neljä epädeterminististä automaattia, jotka ovat $\mathbf{A}_1^{edet}$, $\mathbf{A}_2^{edet}$, $\mathbf{A}_3^{edet}$ ja $\mathbf{A}_4^{edet}$. Näillä automaateilla on muuten sama siirtymäkuvaus kuin automaatilla $\mathbf{A}^{edet}$ paitsi, että poistetaan jokaisesta aiemmin valittu siirtymä $p \xrightarrow{a} q$, mutta säilytetään kaikki automaatin $\mathbf{A}^{edet}$

tilat. Näin ollen automaatit eroavat vain alku- ja lopputilojen osalta:

- Automaatilla $\mathbf{A}_1^{edet}$ on alkutilat I ja lopputilat T
- Automaatilla $\mathbf{A}_2^{edet}$ on alkutilat I ja lopputila $\{p\}$
- Automaatilla $\mathbf{A}_3^{edet}$ on alkutila $\{q\}$ ja lopputila $\{p\}$
- Automaatilla $\mathbf{A}_4^{edet}$ on alkutila $\{q\}$ ja lopputilat T.

Kaikkien näiden automaattien siirtymäluku on $n-1$, joten induktio-oletuksen nojalla kaikki kielet $L(\mathbf{A}_1^{edet})$, $L(\mathbf{A}_2^{edet})$, $L(\mathbf{A}_3^{edet})$ ja $L(\mathbf{A}_4^{edet})$ ovat säännöllisiä.

Osoitetaan, että $L = L_1 + (L_2a)(L_3a)^*L_4$. On helppo osoittaa, että $L_1 \subseteq L$ ja $(L_2a)(L_3a)^*L_4 \subseteq L$, joten keskitytään inkluusioon $L \subseteq L_1 + (L_2a)(L_3a)^*L_4$. Olkoon $x \in L$. Tällöin x on polku automaatissa $\mathbf{A}^{edet}$, joka alkaa jostakin alkutilasta ja päättyy johonkin lopputilaan. Tämä polku joko sisältää aiemmin valitun siirtymän $p \xrightarrow{a} q$ tai sitten se ei sisällä sitä. Jos se ei sisällä sitä, niin tällöin $x \in L(\mathbf{A}_1^{edet})$. Voidaan siis olettaa, että se kuuluu aiemmin mainittuun polkuun. Nyt täytyy etsiä merkin a ne esiintymät merkkijonossa x, jotka vastaavat siirtymään $p \xrightarrow{a} q$. Tällöin merkkijono x voidaan jakaa tekijöihin seuraavasti: $x = (ua)(v_1a)\cdots(v_na)w$, missä u on nimike polulle alkutilasta tilaan p, jokainen merkkijono v_i nimeää polun tilasta q tilaan p ja w on nimike polulle tilasta q lopputilaan. Siten $x \in (L_2a)(L_3a)^*L_4$, jolloin $x \in L_1 + (L_2a)(L_3a)^*L_4$ ja $L = L_1 + (L_2a)(L_3a)^*L_4$. Siis L on säännöllinen. $\square$

3 Büchin automaatit

Edellisessä luvussa tarkasteltiin äärellisiä merkkijonoja. Laajennetaan nyt näkökulmaa äärettömiin merkkijonoihin. Tässä luvussa seurataan teosta [4].

3.1 Äärettömät merkkijonot

Määritelmä 3.1. Ääretöntä jonoa $(a_1, a_2, \ldots, a_n, \ldots)$ aakkoston A merkkejä sanotaan *äärettömäksi merkkijonoksi.* Käytetään merkkijonolle lyhyempää merkintää $a_1 a_2 \cdots a_n \cdots$. A^ω on aakkoston A *kaikkien äärettömien merkkijonojen joukko.* $A^\infty = A^* \cup A^\omega$ on aakkoston A *kaikkien äärellisten tai äärettömien merkkijonojen joukko.* Kuvaus $u : \mathbb{N} \to A$, $u(n) \mapsto a_n$ tuottaa äärettömän merkkijonon. Merkintä $u[r, s]$ tarkoittaa merkkijonoa $u(r)u(r+1) \cdots u(s)$. Merkintää voidaan soveltaa myös äärellisille merkkijonoille.

Huomautus. Merkintä A^ω vastaa merkintää $A^\mathbb{N}$, joka taas on kaikkien kuvauksien joukko luonnollisilta luvuilta aakkostoon A.

Määritelmä 3.2. Olkoon A aakkosto ja olkoot $x \in A^*$ ja $y \in A^\omega$. Tällöin *äärettömän merkkijonon konkatenaatio* on $x \cdot y$. Täytyy huomioida, että $x \cdot y$ on ääretön merkkijono. Merkitään xy merkinnän $x \cdot y$ sijasta.

Esimerkki 3.1. Olkoon $A = \{a, b\}$ aakkosto ja olkoot $x = aa \in A^*$ ja $y = bbb \cdots \in A^\omega$. Tällöin $xy = aabbb \cdots$ ja $xy \in A^\omega$.

Määritelmä 3.3. Olkoot A aakkosto, $x \in A^*$ ja $u \in A^\omega$. Tällöin merkkijono x on äärettömän merkkijonon u *tekijä*, jos on olemassa sellainen merkkijono v ja sellainen ääretön merkkijono w, että $u = vxw$. Merkkijono x on äärettömän merkkijonon u *etuosa*, jos on olemassa sellainen ääretön merkkijono w että $u = xw$.

Esimerkki 3.2. Olkoon $A = \{a, b\}$ aakkosto ja olkoon $ababab \cdots \in A^\omega$. Tällöin äärettömän merkkijonon $ababab \cdots$ etuosia ovat

$$\{\varepsilon, a, ab, aba, abab, ababa, ababab, \ldots\}.$$

Määritelmä 3.4. Olkoot $u, v \in A^\infty$. Laajennetaan määritelmää 2.8, koskemaan myös joukon A^∞ alkioita. Tällöin *puujärjestys yli joukon* A^∞ määritellään asettamalla $u \leq v$, jos $u = v$ tai merkkijono u on äärettömän merkkijonon v etuosa.

Määritelmä 3.5. Kokonaisluku $p > 0$ on äärellisen merkkijonon $u = a_1 a_2 \cdots a_n$ *jakso*, mikäli jokaisella sellaisella kokonaisluvulla k, että $k + p \geq n$, pätee $a_k = a_{k+p}$. Pienintä merkkijonon u jaksoa kutsutaan *merkkijonon u perusjaksoksi.* Myös äärettömän merkkijonon *jakso* ja *perusjakso* määritellään samalla tavalla.

Esimerkki 3.3. Olkoon $A = \{a, b\}$ aakkosto ja $ababbaabbabbba \in A^*$. Tällöin merkkijonon $ababbaabbabbba$ perusjakso on 4, koska

$$ab(abba)^2 bbba = ababbaabbabbba.$$

Määritelmä 3.6. Olkoon A aakkosto. Ääretön merkkijono u on *jaksollinen*, mikäli sen jaksojen joukko on epätyhjä.

Määritelmä 3.7. Kokonaisluku $p > 0$ on äärettömän merkkijonon

$$u = a_1 a_2 \cdots a_n \cdots$$

eräs *ultimaattinen jakso*, mikäli on olemassa sellainen kokonaisluku $k_0 \geq 0$, että jokaisella $k \geq k_0$ pätee $a_k = a_{k+p}$. Jos p ja q ovat merkkijonon u kaksi ultimaattista jaksoa, niin näiden suurin yhteinen tekijä on edelleen äärettömän merkkijonon u ultimaattinen jakso. Pienintä äärettömän merkkijonon u ultimaattista jaksoa kutsutaan *äärettömän merkkijonon u ultimaattiseksi jaksoksi.* Ääretöntä merkkijonoa, jolla on ultimaattinen jakso, kutsutaan *ultimaattisesti jaksolliseksi.*

Määritelmä 3.8. Olkoon A aakkosto. Kaikkien äärettömien merkkijonojen joukossa on määritelty *vaihto* asettamalla kuvaus $\sigma : A^\omega \to A^\omega$ siten, että jokaisella äärettömällä merkkijonolla $u \in A^\omega$ on sellainen kuva $\sigma(u)$, että jokaisella $n \in \mathbb{N}$ pätee $\sigma(u)(n) = u(n+1)$.

Esimerkki 3.4. Olkoon $A = \{a, b\}$ aakkosto ja olkoon $u = abab \cdots$ ääretön merkkijono. Tällöin $\sigma(u) = bab \cdots$.

Määritelmä 3.9. Olkoon L äärettömien merkkijonojen joukko. Tällöin $\sigma(L) = \{\sigma(u) \mid u \in X\}$, missä σ on kuten määritelmässä 3.8. Sanotaan, että L on *vakaa*, mikäli $\sigma(L) \subset L$.

3.2 ω-säännölliset kielet

Laajennetaan määritelmää 2.11, jossa määriteltiin kielten $L, M \subseteq A^*$ tulo.

Määritelmä 3.10. Olkoon A aakkosto, $L \subseteq A^*$ ja $M \subseteq A^\infty$. Tällöin kielien L ja M *tulo* on $L \cdot M = \{xy \mid x \in L, y \in M\}$. Käytetään yleisesti merkintää LM merkinnän $L \cdot M$ sijaan.

Huomautus. Tästä alkaen sovelletaan kielien tulon suhteen määritelmää 3.10.

Lause 3.1. *Olkoot $L, M \subseteq A^*$ ja olkoon $N \subseteq A^\infty$. Tällöin $(LM)N = L(MN)$.*

Todistus. Seuraa suoraan määritelmästä 3.10, joten se sivuutetaan. □

Määritelmä 3.11. Olkoon $L \subseteq A^*$ kieli. Tällöin ω on operaattori, joka määritellään asettamalla $L^\omega = \{x_0x_1\cdots \mid \forall i \geq 0, x_i \in L \setminus \{\varepsilon\}\}$.

Määritelmän 3.11 nojalla kieli L^ω koostuu äärettömistä merkkijonoista, jotka on muodostettu konkatenoimalla äärettömän monta kertaa kielen L epätyhjiä merkkijonoja.

Laajennetaan määritelmää 2.27 siten, että jos (s) on säännöllinen lauseke, niin on myös (s^ω). Jatkossa säännöllisiä lausekkeita kutsutaan *ω-säännöllisiksi lausekkeiksi.*

Esimerkki 3.5. Olkoot $u = a_1a_2\ldots a_n$ ja $L = \{u\}$. Tällöin $L^\omega = \{u^\omega\}$, missä u^ω on ääretön merkkijono

$$u^\omega = a_1a_2\cdots a_na_1a_2\cdots a_na_1a_2\cdots a_na_1a_2\cdots .$$

Lause 3.2. *Olkoot $L, M \subset A^*$ kieliä. Tällöin seuraavat ominaisuudet pätevät:*

i) $(L+M)^\omega = (L^*M)^\omega + (L+M)^*L^\omega$

ii) $(LM)^\omega = L(ML)^\omega$

iii) Jokaisella $n > 0$, $(L^n)^\omega = (L^+)^\omega = L^\omega$

iv) $LL^\omega = L^+L^\omega$.

Todistus. Todistuksissa sovelletaan vain määritelmiä 3.10 ja 3.11, joten ne sivuutetaan. □

Määritelmä 3.12. Olkoon A aakkosto. Tällöin joukon A^ω *ω-säännöllisten kielien luokka* on pienin sellainen joukon A^ω kielien joukko $\mathcal{S}$, että seuraavat ehdot pätevät:

i) $\emptyset \in \mathcal{S}$ ja jokaisella $a \in A$, $\{a\} \in \mathcal{S}$

ii) $\mathcal{S}$ on suljettu äärellisen yhdisteen suhteen

iii) $\mathcal{S}$ on suljettu tulon suhteen

iv) $\mathcal{S}$ on suljettu operaation $*$ suhteen

v) $\mathcal{S}$ on suljettu operaation ω suhteen.

Määritelmän 3.12 nojalla joukon A^∞ ω-säännöllisten kielien luokka $\mathcal{S}$ on pienin joukon A^∞ luokka, joka sisältää kaikki joukosta A^* muodostetut äärelliset kielet ja on suljettu äärellisen yhdisteen, äärellisen tulon, ja operaatioiden $*$ ja ω suhteen.

Lause 3.3. *Olkoot $M, N \subseteq A^*$ säännöllisiä kieliä. Kieli $L \subseteq A^\omega$ on ω-säännöllinen kieli, jos ja vain jos kieli L on muotoa MN^ω olevien säännöllisten kielten äärellinen yhdiste.*

Todistus. Kts. [4, s. 15]. Käytetään väitteessä mainittujen kielten L luokasta merkintää $\mathcal{S}(A^\infty)$. Suunta oikealta vasemmalle seuraa suoraan määritelmästä 3.12.

Osoitetaan suunta vasemmalta oikealle todistamalla yleisempi tulos. Väitetään, että jos $L \subseteq A^\omega$ on säännöllinen kieli, niin tällöin pätee:

i) $L \cap A^*$ on säännöllinen kieli

ii) $L \cap A^\omega \in \mathcal{S}(A^\omega)$.

Riittää osoittaa, että $L \in \mathcal{S}(A^\omega)$, kun $L \subseteq A^\omega$. Olkoon $\mathcal{L}$ kaikkien niiden kielien $L \subseteq A^\infty$ luokka, jotka toteuttavat ehdot (i) ja (ii). Tällöin pätee:

1) $\emptyset \in \mathcal{L}$ ja $\{a\} \in \mathcal{L}$ jokaisella $a \in A$

2) $\mathcal{L}$ on suljettu äärellisen yhdisteen suhteen

3) $\mathcal{L}$ on suljettu tulon suhteen Lisäksi, jos $L \subseteq A^*$ ja $M \subseteq A^\infty$, niin $(LM \cap A^*) = L(M \cap A^*)$, joka on säännöllinen, koska M toteuttaa ehdon (i). Edelleen, jos $(LM \cap A^\omega) = L(M \cap A^\omega)$, joka kuuluu luokkaan $\mathcal{S}(A^\omega)$, koska ehdon (ii) nojalla $M \cap A^\omega \in \mathcal{S}(A^\omega)$.

4) $\mathcal{L}$ on suljettu operaation $*$ suhteen

5) $\mathcal{L}$ on suljettu operaation ω suhteen.

Siispä luokka $\mathcal{L}$ sisältää säännöllisten kielien $L \subseteq A^\infty$ luokan, josta väite seuraa.

□

Esimerkki 3.6. Olkoon $A = \{a, b\}$ aakkosto. Aakkostosta A muodostettujen äärettömien merkkijonojen kieli L, jossa on vain äärellinen määrä merkkejä b, voidaan ilmaista ω-säännöllisillä lausekkeilla $L = (a + b)^* a^\omega$.

3.3 Büchin automaatti

Määritelmä 3.13. *Büchin automaatti* on $\mathbf{A}^b = (S, A, I, E, T)$, missä S on *äärellinen tilojen joukko*, A on *aakkosto*, I on *alkutilojen* joukko, $E \subseteq S \times A \times S$ on *siirtymäkuvaus*, joka vastaa epädeterministisen äärellisen automaatin $\mathbf{A}^{edet}$ siirtymäkuvausta, ja T on lopputilojen joukko. Yleensä Büchin automaatista käytetään merkintää $\mathbf{A}^b$.

Huomautus. Myös äärellisen automaatin tapauksessa voitaisiin tarkastella siirtymäkuvausta vain relaation $S \times A \times S$ osajoukkona. Esimerkiksi mielivaltaisen automaatin siirtymän merkintä $s_0 a = s_1$ tarkoittaa samaa kuin merkintä $e = (s_0, a, s_1)$.

Määritelmä 3.14. Ääretöntä lukumäärää peräkkäisiä siirtymiä

$$e_0 = (s_0, a, s_1), e_1 = (s_1, b, s_2), \ldots$$

missä $e_1, e_2 \in E$, $s_0, s_1, s_2 \in S$ ja $a, b \in A$, kutsutaan *äärettömäksi poluksi.* Merkinnän $e_0 = (s_0, a, s_1), e_1 = (s_1, b, s_2), \ldots$ sijaan merkitään lyhyemmin $e_0 e_1 \cdots$.

Büchin automaatin $\mathbf{A}^b$ siirtymätaulukko ja -verkko ovat samanlaisia kuin epädeterministisen äärellisen automaatin $\mathbf{A}^{edet}$ vastaavat.

Määritelmä 3.15. Olkoon $\mathbf{A}^b = (S, A, I, E, T)$ Büchin automaatti. Ääretöntä polkua kutsutaan *alkuperäiseksi*, mikäli se lähtee jostakin alkutilasta $i \in I$, ja *lopulliseksi*, mikäli se käy lopputilojen joukossa T äärettömän monta kertaa. Ääretöntä polkua kutsutaan *toteutuvaksi*, mikäli se on alkuperäinen ja lopullinen.

Määritelmä 3.16. Automaatin $\mathbf{A}^b = (S, A, I, E, T)$ *tunnistama* kieli on

$$L^\omega(\mathbf{A}^b) = \{x \in A^\omega \mid \mathrm{Inf}(x) \cap T \neq \emptyset\},$$

missä Inf(x) on niiden tilojen joukko, jotka esiintyvät alkuperäisessä äärettömässä polussa äärettömän monta kertaa. Automaatti $\mathbf{A}^b$ siis *hyväksyy* äärettömän merkkijonon $x \in A^\omega$, jos ja vain jos $\mathrm{Inf}(x) \cap T \neq \emptyset$.

Määritelmä 3.17. Kieli $L \subseteq A^\omega$ on *tunnistettavissa*, jos on olemassa sellainen automaatti $\mathbf{A}^b$, että $L = L^\omega(\mathbf{A}^b)$.

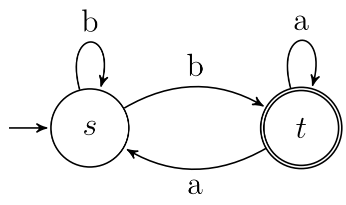
Määritelmä 3.18. Büchin automaatti $\mathbf{A}^b = (S, A, I, E, T)$ on *deterministinen*, mikäli siirtymäkuvaus E on sama kuin automaatin $\mathbf{A}^{det}$ siirtymäkuvaus ja I on yksiö eli alkutiloja on vain yksi. Merkitään determinististä Büchin automaattia $\mathbf{A}^{bdet}$.

Määritelmä 3.19. Automaatti $\mathbf{A}^b$ on *numeroituva*, jos sen tilojen aakkosto on numeroituva.

Esimerkki 3.7. $\mathbf{A}^b = (\{s, t\}, \{a, b\}, s, E, r)$ on Büchin automaatti, jonka siirtymät ovat taulukossa 3.1, ja jonka siirtymäverkko on kuvassa 3.1. Automaatti $\mathbf{A}^b$ tunnistaa kielen $b(b^*a)^\omega$.

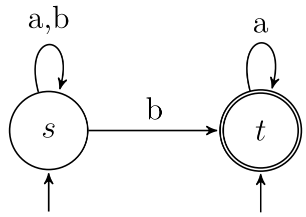
	a	b
$\rightarrow s$		$\{s, t\}$
$\leftarrow t$	$\{s, t\}$	

Taulukko 3.1: Esimerkin 3.7 siirtymätaulukko.



Kuva 3.1: Esimerkin 3.7 siirtymäverkko.

Esimerkki 3.8. $\mathbf{A}^b = (\{s,t\}, \{a,b\}, \{s,t\}, E, t)$ on Büchin automaatti, jonka siirtymät ovat taulukossa 3.2, ja jonka siirtymäverkko on kuvassa 3.2.



Kuva 3.2: Esimerkin 3.8 siirtymäverkko.

	a	b
$\rightarrow s$	s	$\{s,t\}$
$\leftrightarrow t$	t	

Taulukko 3.2: Esimerkin 3.8 siirtymätaulukko.

Lause 3.4. *Jokainen epätyhjä ja tunnistettava kieli $L \subseteq A^\omega$ sisältää ultimaattisesti jaksollisen merkkijonon.*

Todistus. Kts. [4, s. 27]. Olkoon $L \subseteq A^\omega$ epätyhjä ja tunnistettava automaatilla $\mathbf{A}^b = (S, A, I, E, T)$. Koska kieli L on epätyhjä, niin on olemassa polku $p = p_0p_1p_2\cdots$, missä p_0 alkaa alkutilojen joukon I tilasta ja loppuu tilaan $t \in T$, ja missä $p_1p_2\cdots$ ovat polkuja tilasta t tilaan t. Polku $p_0p_1p_1p_1\cdots$ on siten toteutuva polku ja sen nimike on ultimaattisesti jaksollinen merkkijono. □

Määritelmä 3.20. Olkoon $\mathbf{A}^b = (S, A, I, E, T)$ automaatti. Tila $s \in S$ on *saavutettava*, mikäli automaatissa $\mathbf{A}^b$ on olemassa mahdollisesti tyhjä äärellinen alkuperäinen polku, joka loppuu tilaan $s \in S$. Tila $s \in S$ on *kosaavutettava*, mikäli automaatissa $\mathbf{A}^b$ on olemassa ääretön lopullinen polku, joka alkaa tilasta $s \in S$. Automaatti $\mathbf{A}^b$ on *karsittu*, jos kaikki sen tilat saavutettavia ja kosaavutettavia.

Lause 3.5. *Mihin tahansa automaattiin $\mathbf{A}^b$ voidaan liittää sellainen karsittu automaatti $\mathbf{A}^{b'}$, että seuraavat ehdot pätevät:*

i) Automaatit $\mathbf{A}^b$ ja $\mathbf{A}^{b'}$ tunnistavat saman kielen $L \subseteq A^\omega$

ii) Jos $\mathbf{A}^b$ on deterministinen, niin on myös $\mathbf{A}^{b'}$

iii) Jos $\mathbf{A}^b$ on äärellinen, niin on myös $\mathbf{A}^{b'}$.

Todistus. Kts. [4, ss. 27-28]. Tarkastellaan aluksi tapausta (i). Olkoon $\mathbf{A}^b = (S, A, I, E, T)$ automaatti ja olkoon $P \subseteq S$ automaatin $\mathbf{A}^b = (S, A, I, E, T)$ saavutettavien ja kosaavutettavien tilojen joukko. Olkoon $\mathbf{A}^{b'} = (P, A, I \cap P, E', T \cap P)$ automaatti, missä $E' = E \cap (P \times A \times P)$. Selvästi $L^\omega(\mathbf{A}^{b'}) \subseteq L^\omega(\mathbf{A}^b)$. Oletetaan, että $u = a_0 a_1 \cdots \in L^\omega(\mathbf{A}^b)$. On olemassa lopullinen polku

$$p = s_0 \stackrel{a_0}{\to} s_1 \stackrel{a_1}{\to} s_2 \cdots ,$$

jonka nimike on u ja $s_0 \in I$. Tilat $s_0, s_1, \ldots$ ovat saavutettavia ja kosaavutettavia. Näin ollen s on polku automaatissa $\mathbf{A}^{b'}$ ja $u \in L^\omega(\mathbf{A}^b)$. Siispä $L^\omega(\mathbf{A}^b) = L^\omega(\mathbf{A}^{b'})$.

Tarkastellaan sitten tapausta (ii), joka on varsin suoraviivainen. Jos $\mathbf{A}^b$ on deterministinen, niin E' on deterministinen. Jos automaatin $\mathbf{A}^b$ tunnistama kieli on epätyhjä, niin on olemassa lopullinen polku yksikäsitteisestä alkutilasta, joten alkutila on kosaavutettava ja $I \cap P$ on yksiö. Siispä $\mathbf{A}^{b'}$ on deterministinen.

Tapaus (iii) on ilmeinen. □

Büchin automaatin täydellisyys määritellään samoin kuin äärellisen automaatin.

Lause 3.6. *Büchin automaatin $\mathbf{A}^b$ tunnistama kieli $L \subseteq A^\omega$ voidaan tunnistaa sellaisella täydellisellä Büchin automaatilla $L^\omega(\mathbf{A}^{b'})$, jolla pätee: jos $\mathbf{A}^b$ on äärellinen, niin myös $\mathbf{A}^{b'}$ on äärellinen, ja jos $\mathbf{A}^b$ on deterministinen, niin myös $\mathbf{A}^{b'}$ on deterministinen.*

Todistus. Kts. [4, s. 28]. Olkoon $\mathbf{A}^b = (S, A, I, E, T)$ automaatti, joka tunnistaa kielen $L \subseteq A^\omega$.

Jos $\mathbf{A}^b$ ei ole täydellinen, niin olkoon $\mathbf{A}^{b'} = (S \cup \{p\}, A, I, E', T)$, missä p on uusi tila ja $E' = E \cup E_1 \cup E_2$, missä

$$\begin{aligned} E_1 &= \{(p, a, p \mid a \in A\} \\ E_2 &= \{(s, a, p) \mid s \in S, a \in A \text{ ja } (\{s\} \times \{a\} \times S) \cap E = \emptyset\}. \end{aligned}$$

Automaatti $\mathbf{A}^{b'}$ tunnistaa edelleen kielen L ja on täydellinen. Lisäksi sen on deterministinen (vastaavasti äärellinen), jos $\mathbf{A}^b$ on deterministinen (vastaavasti äärellinen). □

3.4 Kleenen lause ω-säännöllisille kielille

Lause 3.7 (Kleenen lause ω-säännöllisille kielille)**.** *Kieli $L \subseteq A^\omega$ on tunnistettava, jos ja vain jos se on ω-säännöllinen.*

Todistus. Kts. [4, ss. 28-29]. Oletetaan, että $X \subseteq A^\omega$ on tunnistettava ja olkoon $\mathbf{A}^b = (S, A, I, E, T)$ äärellinen Büchin automaatti, joka tunnistaa kielen X. Tällöin

$$X = L^\omega(\mathbf{A}^b)^\omega = \bigcup_{i \in I} \bigcup_{t \in T} L^*(S, A, i, E, f)(L^+(S, A, f, E, f))^\omega,$$

joka osoittaa, että kieli X on ω-säännöllinen, koska Kleenen lauseen 2.21 nojalla $L^*(S, A, i, E, f) \subseteq A^*$ ja $L^+(S, A, f, E, f) \subseteq A^*$ ovat säännöllisiä kieliä.

Oletetaan sitten, että $Y \subseteq A^\omega$ on ω-säännöllinen kieli $X(X')^\omega$, missä $X, X' \subseteq A^*$ ovat säännöllisiä kieliä. Olkoot $\mathbf{A} = (S, A, i, E, t)$ ja $\mathbf{A}' = (S, A, i', E, t')$ sellaisia normalisoituja [3, s. 109] automaatteja, että $X = L^+(\mathbf{A})$ ja $X' = L^+(\mathbf{A}')$. Nyt $\mathbf{A}^B = ((S \cup S') \setminus \{i', t'\}, i, E^\#, t)$, missä $E^\# = E \cup E_0 \cup E_1 \cup E_2$, kun

$$\begin{aligned} E_0 &= \{(t, a, t) \mid (i', a, t') \in E'\} \\ E_1 &= \{(t, a, s) \mid s \in S' \setminus \{i', t'\} \text{ ja } (i', a, s) \in E'\} \\ E_2 &= \{(s, a, t) \mid s \in S' \setminus \{i', t'\} \text{ ja } (s, a, t') \in E'\}. \end{aligned}$$

Näin ollen kieli $X(X')^\omega$ on tunnistettavissa. Osoitetaan vielä, että tunnistettavien kielien $L \subseteq A^\omega$ luokka on suljettu äärellisen yhdisteen suhteen.

Olkoot $Y, Y' \subseteq A^\omega$ tunnistettavia kieliä, jotka ovat tunnistettavissa äärellisillä Büchin automaateilla $\mathbf{A}^b = (S, A, I, E, T)$ ja $\mathbf{A}^{b'} = (S', A, I', E', T')$. Voidaan olettaa, että S ja S' ovat erilliset ja siten voidaan todeta, että siirtymäkuvaukset E ja E' ovat relaation $(S \cup S') \times A \times (S \cup S')$ osajoukkoja. Täten seuraava identiteetti pätee

$$Y \cup Y' = L^\omega(\mathbf{A}^b) \cup L^\omega(\mathbf{A}^{b'}) = L^\omega(S \cup S', A, I \cup I', E \cup E', T \cup T')$$

ja siten $Y \cup Y'$ on tunnistettavissa äärellisellä automaatilla $(S \cup S', A, I \cup I', E \cup E', T \cup T')$. □

Määritellään kielille uusi operaattori $\rightarrow$ ja tarkastellaan sen toimintaa esimerkin kautta.

Määritelmä 3.21. Olkoon $L \subseteq A^*$ kieli. Tällöin

$$\overrightarrow{L} = \{u \in A^\omega \mid \text{merkkijonolla } u \text{ on äärettömästi etuosia kielessä } L\},$$

on myös kieli.

Esimerkki 3.9. Jos $L_1 = a^*b$, niin $\overrightarrow{L_1} = \emptyset$. Jos $L_2 = (ab)^+$, niin $\overrightarrow{L_2} = (ab)^\omega$. Jos $L_3 = (a^*b)^+ = (a+b)^*b$, niin $\overrightarrow{L_3} = (a^*b)^\omega$.

Esimerkki 3.10. Kieltä $X = (a+b)^*a^\omega$, jossa merkkejä b on äärellinen määrä, ei saada mistään kielestä L operaatiolla $\overrightarrow{L}$. Jos se saataisiin, niin tällöin merkkijonolla ba^ω olisi etuosa $u_1 = ba^{n_1} \in L$, merkkijonolla $ba^{n_1}ba^{n_2}$ olisi etuosa $u_2 = ba^{n_1}ba^{n_2} \in L, \ldots$, ja äärettömällä merkkijonolla $u = ba^{n_1}ba^{n_2}ba^{n_3}\cdots$ olisi äärettömästi etuosia kielessä L. Tällöin $u \in \overrightarrow{L}$, mutta tämä on mahdotonta, koska u sisältää äärettömästi merkkejä b.

Lause 3.8. *Olkoon $\mathbf{A}^{bdet}$ automaatti. Tällöin $L^\omega(\mathbf{A}^{bdet}) = \overrightarrow{L^+(\mathbf{A}^{bdet})}$.*

Todistus. Kts. [4, s. 32]. Olkoon $\mathbf{A}^{bdet} = (S, A, i, E, T)$. Jos $u \in L^\omega(\mathbf{A}^{bdet})$, niin tällöin ääretön merkkijono u on polun

$$p = (s_0, a_0, s_1)(s_1, a_1 s_2)\cdots$$

sellainen nimike, että $s_0 = i$ ja on olemassa sellainen lukujono $n_0 < n_1 < \cdots$, että $s_{n_0}, s_{n_1}, \ldots \in T$. Merkkijonot $u_k = a_0 a_1 \cdots a_{n_{k-1}} \in L^+(\mathbf{A}^{bdet})$ ja ovat äärettömän merkkijonon u etuosia. Siispä $L^\omega(\mathbf{A}^{bdet}) \subseteq \overrightarrow{L^+(\mathbf{A}^{bdet})}$.

Jos $u \in \overrightarrow{L^+(\mathbf{A}^{bdet})}$, niin merkkijonolla u on äärettömästi etuosia kielessä $L^+(\mathbf{A}^{bdet})$. Koska $\mathbf{A}^{bdet}$ on deterministinen, niin u on alkuperäisen polun nimike, joka käy äärettömästi joukon T tiloissa. Siten $u \in L^\omega(\mathbf{A}^{bdet})$. Näin on osoitettu, että $L^\omega(\mathbf{A}^{bdet}) = \overrightarrow{L^+(\mathbf{A}^{bdet})}$. □

Lause 3.9. *Olkoon $X \subseteq A^\omega = (S, A, I, E, T)$ kieli. Tällöin seuraavat ehdot ovat ekvivalentteja:*

i) Kieli X voidaan tunnistaa automaatilla $\mathbf{A}^{bdet}$.

ii) On olemassa sellainen kieli $L \subseteq A^+$, että $X = \overrightarrow{L}$.

Lisäksi, jos aakkosto A on numeroituva, niin ehdot (i) ja (ii) ovat ekvivalentteja ehdon (iii) kanssa.

iii) X voidaan tunnistaa numeroituvalla automaatilla $\mathbf{A}^{bdet}$.

Todistus. Kts. [4, ss. 32-33]. Jos $\mathbf{A}^{bdet}$ on automaatti, joka tunnistaa kielen X, niin $X = L^\omega(\mathbf{A}^{bdet}) = \overrightarrow{L^+(\mathbf{A}^{bdet})}$ lauseen 3.8 nojalla. Näin ollen (i) implikoi (ii).

Olkoon sitten $L \subseteq A^+$ sellainen, että $X = \overrightarrow{L}$. Tällöin L on tunnistettavissa deterministisellä automaatilla $\mathbf{A}^{bdet} = (A^*, A, \varepsilon, \cdot, L)$, missä $\cdot$ on deterministinen siirtymäkuvaus, joka on määritelty jokaisella $u \in A^*$ ja jokaisella $a \in A$ asettamalla $u \cdot a = ua$. Lauseen 3.8 perusteella $L^\omega(\mathbf{A}^{bdet}) = \overrightarrow{L^+(\mathbf{A}^{bdet})} = \overrightarrow{L} = X$. Siten (ii) implikoi (i).

Jos aakkosto A numeroituva, niin $\mathbf{A}^{bdet}$ on numeroituva ja tällöin ehdosta (ii) seuraa ehto (iii). Koska ehdosta (iii) seuraa selvästikin ehto (i), niin kaikkien kolmen ehdon välille muodostuu ekvivalenssi. □

Lause 3.10. *Olkoon $X \subseteq A^\omega$. Tällöin seuraavat ehdot ovat ekvivalentteja:*

i) Kieli X voidaan tunnistaa äärellisellä Büchin automaatilla $\mathbf{A}^{bdet}$.

ii) On olemassa sellainen tunnistettava kieli $L \subseteq A^+$, että $X = \overrightarrow{L}$.

Todistus. Kts. [4, s. 33]. Lauseen 3.9 todistusta voidaan soveltaa äärelliseen tapaukseen. Jos $\mathbf{A}^{bdet}$ on äärellinen, niin tällöin $L^+(\mathbf{A}^{bdet})$ on tunnistettavissa. Siis ehdosta (i) seuraa ehto (ii). Toiseen suuntaan riittää valita sellainen äärellinen $\mathbf{A}^{bdet}$, joka tunnistaa kielen L. □

Seuraavassa lauseessa käytetään poikkeuksellisesti tilojen joukolle merkintää Q.

Lause 3.11. *Mikä tahansa determinististen kielten äärellinen yhdiste (vastaavasti äärellinen leikkaus) on deterministinen.*

Todistus. Kts. [4, ss. 33-34]. Äärellisen yhdisteen tapaus seuraa kaavasta

$$\overrightarrow{\bigcup_{1 \leq i \leq n} L_i} = \bigcup_{1 \leq i \leq n} \overrightarrow{L_i}$$

ja lauseesta 3.9.

Leikkausta varten tarvitaan seuraava konstruktio. Tarkastellaan kahta determinististä Büchin automaattia $\mathbf{A_1}^{bdet} = (Q_1, A, i_1, E_1, T_1)$ ja $\mathbf{A_2}^{bdet} = (Q_2, A, i_2, E_2, T_2)$. Olkoon

$$Q = Q_1 \times Q_2 \times \{0, 1\}$$

ja olkoon $\mathbf{A}^{bdet} = (Q, A, \{i_1, i_2, 1\}, E, T)$ deterministinen automaatti, missä

$$\begin{aligned} E = &\{((p_1, p_2, s), a, (q_1, q_2, t)) \mid (p_1, a, q_1) \in E_1, (p_2, a, q_2) \in E_2 \text{ ja} \\ &\quad t = 0 \text{ jos ja vain jos } (s = 1, p_2 \in T_2 \text{ ja } q_1 \notin T_1) \text{ tai } (s = 0 \text{ ja } q_1 \notin T_1)\} \end{aligned}$$

ja

$$T = Q_1 \times T_2 \times \{1\}.$$

Tilan (q_1, q_2, s) komponentin s tehtävänä on muistaa, että nykyinen polku on kulkenut joukon T_1 tilan kautta viimeisen joukossa T_2 käynnin jälkeen. Jos

$$(p_1, p_2, s) \xrightarrow{a} (q_1, q_2, t)$$

on siirtymä, ja jos $p_2 \in T_2$, niin pätee

$$t = \begin{cases} 1, \text{ jos } q_1 \in T_1 \\ 0, \text{ jos } q_1 \notin T_1. \end{cases}$$

Mutta, jos $p_2 \notin T_2$, niin

$$t = \begin{cases} 1 & , \text{ jos } s = 1 \text{ tai jos } q_1 \in T_1 \\ 0 & , \text{ jos } s = 0 \text{ ja jos } q_1 \notin T_1. \end{cases}$$

Osoitetaan, että $L^\omega(\mathbf{A}^{bdet}) = L^\omega(\mathbf{A_1}^{bdet}) \cap L^\omega(\mathbf{A_2}^{bdet})$. Olkoon

$$c = (q_{0,1}, q_{0,2}, s_0) \xrightarrow{a_0} (q_{1,1}, q_{1,2}, s_1) \xrightarrow{a_1} (q_{2,1}, q_{2,2}, s_2) \cdots$$

polku automaatissa $\mathbf{A}^{bdet}$, jonka nimike on $u = a_0a_1a_2 \cdots$. Polun c avulla saadaan määriteltyä sellainen automaatin $\mathbf{A_1}^{bdet}$ polku c_1 ja sellainen automaatin $\mathbf{A_2}^{bdet}$ polku c_2, että

$$\begin{aligned} c_1 &= q_{0,1} \xrightarrow{a_0} q_{1,1} \xrightarrow{a_1} q_{2,1} \cdots \\ c_2 &= q_{0,2} \xrightarrow{a_0} q_{1,2} \xrightarrow{a_1} q_{2,2} \cdots . \end{aligned}$$

Oletetaan, että c on automaatissa $\mathbf{A}^{bdet}$ toteutuva polku. Tällöin $q_{0,1} = i_1$, $q_{0,2} = i_2$ ja $s_0 = 1$. Polku c kulkee äärettömästi joukon T tilojen kautta, joten on olemassa sellainen ääretön lukujono $(n_k)_{k \geq 0}$, että $q_{n_k,2} \in T_2$ ja $s_{n_k} = 1$. Välittömänä seurauksena polku c_2 on automaatissa $\mathbf{A_2}^{bdet}$ toteutuva polku. Myös polku c_1 on toteutuva, koska $s_{n_k} = 1$ nojalla viimeisimmän joukon T_2 jonkin tilassa käynnin jälkeen on vierailtu ainakin yhdessä joukon T_1 tilassa. Siispä $u \in L^\omega(\mathbf{A_1}^{bdet}) \cap L^\omega(\mathbf{A_2}^{bdet})$. Oletetaan sitten, että $u \in L^\omega(\mathbf{A_1}^{bdet}) \cap L^\omega(\mathbf{A_2}^{bdet})$. Tällöin u on automaatissa $\mathbf{A_1}^{bdet}$ toteutuvan polun c_1 nimike ja automaatissa $\mathbf{A_2}^{bdet}$ toteutuvan polun c_2 nimike:

$$\begin{aligned} c_1 &= q_{0,1} \xrightarrow{a_0} q_{1,1} \xrightarrow{a_1} q_{2,1} \cdots \\ c_2 &= q_{0,2} \xrightarrow{a_0} q_{1,2} \xrightarrow{a_1} q_{2,2} \cdots . \end{aligned}$$

Ääretön merkkijono u määrittelee siis alkuperäisen polun c automaatissa $\mathbf{A}^{bdet}$. Jos $q_{n,2} \in T_2$, niin $q_{n',1} \in T_1$ jollakin $n' \geq n$. Jos $n'' > n'$ on pienin sellainen kokonaisluku, että kun $q_{n'',2} \in T_2$, niin $(i_1, i_2, 1) \cdot a_0a_1 \cdots a_{n''} = (q_{n'',1}, q_{n'',2}, 1) \in T$, jolloin c on toteutuva. Näin ollen $u \in L^\omega(\mathbf{A}^{bdet})$. $\square$

Deterministisellä äärellisellä automaatilla ja epädeterministisellä äärellisellä automaatilla on sama ilmaisuvoima kuten luvussa 2 osoitettiin. Näin ei kuitenkaan ole deterministisen Büchin automaatin ja epädeterministisen Büchin automaatin tapauksessa. Toisin sanoen determinisointi ei onnistu. Esimerkissä 3.10 osoitettiin, että kieli $X = (a + b)^*a^\omega$ ei ole muotoa $\overrightarrow{L}$. Mutta kuitenkin kieli $X = (a + b)^*a^\omega$ voidaan tunnistaa esimerkin 3.8 automaatilla. Siis on olemassa kieliä, joka voidaan tunnistaa epädeterministisellä Büchin automaatilla, mutta ei deterministisellä Büchin automaatilla.

Lähteet

[1] Hopcroft, J. E., Ullman, J. D. *Introduction to Automata Theory, Languages and Computation*. Menlo Park: Addison-Wesley Publishing Company, 1979.

[2] Howie, J. M. *Automata and Languages*. New York: Oxford University Press, 1991.

[3] Lawson, M. V. *Finite Automata*. Boca Ranton: CRC Press LLC, 2004.

[4] Perrin, D., Pin, J-E. *Infinite Words: Automata, Semigroups, Logic and Games*. Amsterdam: Elsevier, 2004.

[5] Rozier, K. Y. *Survey: Linear Temporal Logic Symbolic Model Checking*. Amsterdam: Elsevier, 2011.